

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2002, 00:38:12 ; Search time 1750.92 Seconds
(without alignments)
5858.455 Million cell updates/sec

Title: us-09-821-821-1
Perfect score: 760
Sequence: 1 ttccagtgctccagcagcc.....aaaaaaaaaaaaaaaaaaaaa 760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

- 1: em_estb.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estopl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_estl.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 707.2 | 93.1 | 713 | 10 | BI829760 |
| 2 | 504.8 | 66.4 | 516 | 9 | AA436088 zu03a08.r |
| 3 | 503.2 | 66.2 | 508 | 9 | AI149899 qf43r06.x |
| 4 | 382.2 | 50.3 | 387 | 9 | AA416972 |
| 5 | 377.2 | 49.6 | 382 | 9 | AA435988 zu03a08.s |
| 6 | 325.4 | 42.8 | 394 | 9 | AA435988 zu03a08.s |
| 7 | 300.4 | 39.5 | 415 | 9 | AA470059 zu03a08.s |
| 8 | 295.2 | 38.8 | 484 | 9 | BE107659 ut-r-BT1- |
| 9 | 271 | 35.7 | 482 | 9 | AI002083 |
| 10 | 267.4 | 35.2 | 410 | 9 | AA411806 |
| 11 | 245.4 | 32.3 | 389 | 9 | AA781801 |
| 12 | 148.4 | 19.5 | 538 | 12 | AQ108532 CIT-HSP-2 |
| 13 | 145 | 19.1 | 281 | 9 | AA707529 ab41a12.s |
| 14 | 144.2 | 19.0 | 332 | 10 | BF319786 |
| 15 | 135 | 17.8 | 601 | 10 | BE638317 |
| 16 | 124.6 | 16.4 | 224 | 10 | BE638325 |
| 17 | 95 | 12.5 | 568 | 10 | BM253528 |

| | | | | | |
|----|------|------|------|----|----------|
| 18 | 86.4 | 11.4 | 790 | 10 | BI771845 |
| 19 | 86.4 | 11.4 | 1071 | 9 | AL544561 |
| 20 | 85 | 11.2 | 900 | 10 | BI759693 |
| 21 | 84 | 11.1 | 793 | 10 | BG484817 |
| 22 | 82.8 | 10.9 | 799 | 10 | BI908709 |
| 23 | 82.2 | 10.8 | 704 | 10 | BG571626 |
| 24 | 82 | 10.8 | 857 | 10 | BG720182 |
| 25 | 81.4 | 10.7 | 218 | 10 | BE638321 |
| c | 80.4 | 10.6 | 508 | 9 | AA234138 |
| 26 | 80.4 | 10.6 | 508 | 9 | AA234138 |
| 27 | 80.4 | 10.6 | 562 | 9 | AA418443 |
| 28 | 79.2 | 10.4 | 517 | 10 | BG664654 |
| 29 | 76.2 | 10.0 | 713 | 10 | BI562282 |
| 30 | 72.2 | 9.5 | 421 | 12 | B86842 |
| 31 | 70.8 | 9.3 | 899 | 10 | BG720623 |
| 32 | 68.6 | 9.0 | 921 | 10 | BI561611 |
| 33 | 68.2 | 9.0 | 856 | 10 | BI560592 |
| 34 | 67.6 | 8.9 | 775 | 10 | BI560610 |
| 35 | 66.8 | 8.8 | 689 | 10 | BG719295 |
| 36 | 59.6 | 7.8 | 887 | 9 | AL531049 |
| 37 | 59.4 | 7.8 | 715 | 10 | BG538851 |
| 38 | 58.8 | 7.7 | 579 | 10 | BE513276 |
| c | 58.6 | 7.7 | 470 | 9 | AI807884 |
| 40 | 58.2 | 7.7 | 528 | 10 | BE675149 |
| c | 58.2 | 7.7 | 551 | 10 | BM152988 |
| 42 | 58.2 | 7.7 | 904 | 10 | BG024663 |
| 43 | 58 | 7.6 | 770 | 10 | BI464864 |
| 44 | 57.2 | 7.5 | 947 | 11 | AK008652 |
| 45 | 56 | 7.4 | 1504 | 11 | AK003110 |

ALIGNMENTS

RESULT 1

BI829760
LOCUS 603079831F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171526 5', mRNA linear EST 04-OCT-2001
DEFINITION mRNA sequence.
ACCESSION BI829760
VERSION BI829760.1 GI:15941310
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 713)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>

Plate: LLM11426 row: n column: 07
High quality sequence stop: 713.
Location/Qualifiers
1. .713
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171526"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="PH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range

FEATURES
source

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 202 a 154 c 123 g 233 t 1 others

ORIGIN

Query Match 93.1%; Score 707.2; DB 10; Length 713;
Best Local Similarity 99.4%; Pred. No. 9.7e-90;
Matches 709; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 caggcagctcagcacaagaaagacatggtctagactgaagtaccactaaatcatct 71
|||||
DB 1 CAGGAGCCTCAGCACAGAAAGAACATGGTCTAGACTGAAGTACCACTAAATCATCT 60
|||||
QY 72 cctttcaaatattaccagcaccatcatgattcaagcaccgacacagtcoggtgttcc 131
|||||
DB 61 CCTTCAAAATTATCAGGACACCATCATGGATTCAGACCCGACACAGTCGGGTGTTTC 120
|||||
QY 132 tggattctccagaatacactgcttcagaatatgagtcacagaaactttcagccacga 191
|||||
DB 121 TGGTATTTCCTCCAGAAATCAGTCTTCAGATATGATGATCCAGAACTTTCAGCCACGA 180
|||||
QY 192 cctttcaactcaagcccttcgaaatatttctagagaaatgaaatcttaggga 251
|||||
DB 181 CCTTTCACTCAAGCCCTTGCAAAATATTGCTGAAAAATGAAATCTTAGGGA 240
|||||
QY 252 ctatccagatcctgttgggaattatgacctttctttgaggtatctctctttcaactt 311
|||||
DB 241 CTATCCAGATCCGTGTTGGAAATATGACCTTTCCTTTGAGGTATCTCTCTTTTCACCT 300
|||||
QY 312 tgttaaacacatccaaggtttccctttatattttctcaggatattccattctggggt 371
|||||
DB 301 TGTAAAAACCATATCCAAAGGTTTCCCTTTATATTCTTTCAGGATATCCATTCCTGGGCT 360
|||||
QY 372 ctgtttgttcattaatctcggagccttcctaattcagtgaaagaaagaaacacagaaa 431
|||||
DB 361 CTGTTTNGTCATTAATCTCGAGCCTTCCTAATTCAGTGAAGAAAGAAACACAGAAA 420
|||||
QY 432 ctctgataatattgagccggaataatgaattttctagtcctcgggagcaatagctggaa 491
|||||
DB 421 CTCGATAAATATGAGCGGAATAATGAATTTCTTAGTCCCTGCGGAGCAATAGCTGGA 480
|||||
QY 492 tcattctctcacattggtttctatcctcagatcaaaactacatttgggttatctcacc 551
|||||
DB 481 TCATTCTCCTCACATTGGTTTCATCCTAGATCAAAACTACATTTGTGGTTATCTCACC 540
|||||
QY 552 aaaaatgctagtgtaagcgttactgtctgtcttgggaattttgattacattgatga 611
|||||
DB 541 AAATAGTCAGTGTAGGCTGTACTGTCTCTGTCTGTGGGAATTTGATACATTGATGA 600
|||||
QY 612 ctttcagcattattgaattattcattctcgtccttctcaattttgggggtgacctcag 671
|||||
DB 601 CTTTCACCATTAATGAATTAATTCATTTCTGCTTTCATTTTCGGGTGCCACTCAG 660
|||||
QY 672 aggattgtatgtgaacaatttgttggactagcactgtgagataaaagatgt 724
|||||
DB 661 AGGATTGTGTTGTGAACAATGCTGTGACTAGTGTGAGATAAAGATGT 713
|||||

RESULT 2

AA436088
LOCUS z003a08.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 5' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.

ACCESSION AA436088
VERSION AA436088.1 GI:2141002
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 516)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
JOURNAL
COMMENT Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 733 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 482.

FEATURES
source
1..516
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone="IMAGE:730742"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGGGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 149 a 117 c 80 g 170 t

ORIGIN

Query Match 66.4%; Score 504.8; DB 9; Length 516;
Best Local Similarity 99.6%; Pred. No. 1.7e-61;
Matches 506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 47 gactgaagtaccacaataatcatctccttcaaaattatccaccagacacatcatgattca 106
|||||
DB 1 GACTGAAGTACCACAACTAAGTCACTCTCTTCAAAATTATCACCAGACCATCATCGATTCA 60
|||||
QY 107 agcaccgcacagtcgcggtgtttctggtatttctccacagaaatcactgcttcagaatat 166
|||||
DB 61 AGCACCAGCACAGTCCGGTGTTCCTGGTATTCCTCCAGAAATCACTGCTTCAGAAATAT 120
|||||
QY 167 gagtccacagaaacttccagcagaccttccaaactcaaaagcccttcgaaaaattattt 226
|||||
DB 121 GAGTCCACAGAACTTTCAGCCAGCACCTTTTCAACTCAAAAGCCCTTTCGCAAAATATTT 180
|||||
QY 227 gctagaaaaatgaaaaatcttaggactatccagatccctgtttggaattatgacctttct 286
|||||
DB 181 GCTAGAAAAATGAAATCTTAGGACATATCCAGATCCCTGTTGGAATATATGACCTTTCT 240
|||||
QY 287 ttggagttattctctcttcttcaactttgttaaaacatatccaaaggtttccctttatatt 346
|||||
DB 241 TTTGGAGTTATCTTCTCTTTTTCACCTTTGTTAAACCATATCCAAAGTTTCCCTTTATATTT 300
|||||
QY 347 ctttcagatatccattctggggtctgttttctgttcatttaattcttgaggccttccaat 406
|||||
DB 301 CTTTCAGGATATCCATTCTGGGGCTCTCTTTTGTTCATTAATTCTGGAGCCTTCTTAAT 360
|||||
QY 407 gcagtgaagaaaaaacacagaaactctgataattattgagccgaataaataaatttctt 466
|||||
DB 361 GCAGTGAAGAAAAAACCCACAGAAACTCTGATATATATGACCCGAATATGAAATTTCTT 420
|||||

Seq primer: -41ml3 fwd, ET from Amersham
High quality sequence stop: 280.
Location/Qualifiers
1. .382
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone_lib="IMAGE:730742"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dt) primer [5',
TGTTACCAATCTGAATGGAGCGCGCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

Query Match 50.3%; Score 382.2; DB 9; Length 387;

Best Local Similarity 99.2%; Pred. No. 2.4e-44;

Matches 384; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 346 tcttcaggatccattctgggctctgtttgttcatttaatttgagccttcctaat 405

Db 387 TCTTAGGGATATCCATCTCGGGCTCTGTTTGTTCATTAATCTGAGCCTTCTAAT 328

Qy 406 tgcagtgaaagaagaacacagaaactctgataatattgagccgaataagaattttct 465

Db 327 TGCAGTGAAGAAGAAACACAGAAACTCTGATAATATTGAGCCGAATAATGAATTCT 268

Qy 466 tagtgcctggagcaatagctgaatcattctccacattgtttcattccctagatca 525

Db 267 TAGTGCCTGGAGCAATAGCTGAATCATCTCTCACAATTTGTTTCATCTTAGATCA 208

Qy 526 aaactacattgtgttattctcaccataatagtcagtgataagctgttactgtctgtt 585

Db 207 AACATCAATTTGTGTTTCTCCCAAAATGCTAGTGAGGCTGTACTGCTTCTGTT 148

Qy 586 ctgggaatttgattacattgatgacttccagcattattgaattcattctctctgcc 645

Db 147 CTGGAATTTTGTATTCATATGATGACTTTCAGCAATTAATTAATTCATTTCTCTGCC 88

Qy 646 ttctcaatttgggtggccactcaagatgtgaattgtgaacaattgtttgactagc 705

Db 87 TTTCTCAATTTTGGGGTCCACACACAGGATGTGATGTGGAACAATGTTGTTGACTAGC 28

Qy 706 actgtgagaataaagatgtgttaaat 732

Db 27 ACTGTGAGAATAAAGATGTGTTAAAT 1

RESULT 5
AA435988/c
LOCUS
DEFINITION
3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AA435988.1 GI:2140902
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 382)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK

TITLE
JOURNAL
COMMENT
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevaton.wustl.edu
This clone is available royalty-free through LNL : contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone; similarity on wrong strand
Insert length: 733 Std Error: 0.00

Seq primer: -41ml3 fwd, ET from Amersham
High quality sequence stop: 280.
Location/Qualifiers
1. .382
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone_lib="IMAGE:730742"
/sex="male"
/lab_host="DH10B"

FEATURES

source
Location/Qualifiers
1. .382
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone_lib="IMAGE:730742"
/sex="male"
/lab_host="DH10B"

Query Match 49.6%; Score 377.2; DB 9; Length 382;

Best Local Similarity 99.2%; Pred. No. 1.2e-43;

Matches 379; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 356 tatccattctgggctctgtttgttcatttaattctggagccttcctaatgagtgaaa 415

Db 382 TATCCATCTCGGGCTCTGTTTGTTCATTAATTTGGAGCCCTCTTAATTTGACAGTAAA 323

Qy 416 aaaaaaacacagaaactctgataatattgagccgaataaattttcttagtgcctg 475

Db 322 AGAAAAACACAGAACTCTGATAATATTGACCGCAATAATTAATTTCTTAGTGCCTG 263

Qy 476 ggagcaatagctgggaatcattctccacattgtgttccattcctagatcaaaactacatt 535

Db 262 GGAGCAATAGCTGGAATCATTTCTCCACATTTGTTTTCATCTAGATCAAAATCATTT 203

Qy 536 tgtgttattctcaccataatagtcagtgataagctgttactgcctgttcttgggaatt 595

Db 202 TGTGTTATTCTCACCAAAATAGTCAGTCTAAGGCTGTGTACTGTCTCTTCTGGGAAT 143

Qy 596 ttgattacattgatgacttccagcattattgaattcattctctgccttctcatt 655

Db 142 TTGATTACATGATGATGATTTTCACATTTATTGAATTAATTCATTTCTGCTGCTCAATT 83

Qy 656 ttgggtgcccactcagagattgtgatttgaaacattgtgtgactagcactgtgaaa 715

Db 82 TTGGGTGCCACTCAGAGGATTCTGATTGTGAACAATGTTGTTGACTAGCACTGTGAGAA 23

Qy 716 aaagatgtgttaaatcctcaa 737

Db 22 TAAAGATGTGTTAAATATATAA 1

RESULT 6
AA758635/c
LOCUS
DEFINITION
394 bp mRNA linear EST 29-DEC-1998
ab67b04.s1 Soares testis_NHT Homo sapiens cDNA clone 1320655 3',
similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AA758635.1 GI:2806498
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 394)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

| | |
|---------|---|
| TITLE | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index |
| JOURNAL | Unpublished (1997) |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Ronaldo, Ph.D. |

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 661 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 390.

```

FEATURES
source
n1n1 query sequence: 50bp: 350.
location/Qualifiers
i. .394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1320655"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'.
TGTTCACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pMT3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bepito Soares and M. Fatima Ronaldo.
```

| RESULT | 7 | | | | | |
|-------------|---|--|--|--|--|--|
| AA470059 | | | | | | |
| LOCUS | | | | | | |
| AA470059 | | | | | | |
| 415 bp | | | | | | |
| mRNA | | | | | | |
| linear | | | | | | |
| EST | | | | | | |
| 09-NOV-1997 | | | | | | |

```

DEFINITION      zt94h05.rl Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041
5' similar to SW:CD20_HUMAN F11836 B-LIMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION       AA470059
VERSION         AA470059.1
KEYWORDS        GI:2197368
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE       1 (bases 1 to 415)
AUTHORS         Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Krisman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
WASHU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 656 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop to 393.

```

```

FEATURES
source
    Location/Qualifiers
        i..415
            /organism="Homo sapiens"
            /db_xref="GDB:5926570"
            /db_xref="taxon:9606"
            /clone="IMAGE:730041"
            /clone_lib="Soares_testis_NHT"
            /sex="male"
            /lab_host="DH10B"
            /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TCTTACCAATCTGAAGTGGGCGGCGCCAAATTTTTTTTTTTT 3'],
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
129 a 96 c 67 g 123 t
BASE COUNT
ORIGIN

```

| | Best Local Similarity | 79.9%; | Pred. No. 5.5e-33; | |
|----|-----------------------|---|--------------------|---------------|
| | Matches | 414; | Conservative | 0; Mismatches |
| | | | | 1; Indels |
| | | | | 103; Gaps |
| | | | | 1; |
| Qy | 42 | gtctagactgaagtcaccaataatcatctcttttcaaatattcacgcgacccatcatgg | 101 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 1 | GACTAGACTCAAGTACCAACTCAATCATCTCTTTCAAATTTATCACCGACACCATCATGG | 60 | |
| Qy | 102 | attcaaacacgcgcacacagctccgggtgtttcttggtatttcttccagaataatcaactgcttcag | 161 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 61 | ATTCAAGCACCGCACACAGTCCGGGTGTTCTGGTATTTCCTCCAGAAATCAC7GCTTCAG | 120 | |
| Qy | 162 | aatatgagtccacagaaactttcagccagacacttttcactcaagcccttgcacaaat | 221 | |
| | | | | |
| | | | | |
| Db | 121 | AATATGAGTCCACAGAACTTTCAGCCACGACCTTTTCAACTCAAAGCCCTTGCAAAAT | 180 | |
| Qy | 222 | tatttgc tagaaaaatgaaaactcttagggactatccagatcctgtttggaaattatgacct | 281 | |
| | | | | |
| | | | | |
| Db | 181 | TATTGCTAGAAAATGAAATCTTAGG----- | 208 | |
| Qy | 282 | ttctctttggagttatctctcttttcaacttgttaaacacattcccaaggtttcccttta | 341 | |

Qy 282 ttctctttggagttatcttcccttttccactttgttaaaaccatataccaaaggtttcccttta 341

```

Db 209 ----- 208
QY 342 tattcttcaggatataccattctgggctctgttttggcttatttgcatttgcagcctcc 401
Db 209 -----GGATATCCATTCCTGGGCTCTGTTTGTGTTATTAATTCGGAGCCTCC 257
QY 402 taatlgcagtgaaagaaacacacaaactctgataattgagcgaataatgaatt 461
Db 258 TAATTCAGCTGNAAGAAACACAAACTCTGATATATATTGAGCGAATAATGAAT 317
QY 462 ttcttagtgccctgggagcaatagctgggaatcattctctcatttgcatttgcctcctag 521
Db 318 TTCTTAGTGCCCTGGGAGCAATAGCTGGAATCATCTCTCATATTTGGTTTCATCCTAG 377
QY 522 atcaaaactacatttgggttattctcaccaaaatagt 559
Db 378 ATCAAAACTACTATTGTGTGTTATTCACCAAAATAGT 415

RESULT 8
BE107659/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
484 bp mRNA linear EST 13-JUN-2000
UI-R-BT1-ame-d-04-0-UI.s1 UI-R-BT1 Rattus norvegicus cDNA clone
UI-R-BT1-ame-d-04-0-UI 3', mRNA sequence.
BE107659
BE107659.1 GI:8495769
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 484)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized testis library cDNA Library Preparation: M.B. Soares Lab
Clone distribution: clones will be available through Research
Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-yes.
Location/Qualifiers
1. .484
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db.xref="taxon:10116"
/clone="UI-R-BT1-ame-d-04-0-UI"
/clone_lib="UI-R-BT1"
/dev.stage="adult"
/lab.host="D4108 (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; The library
UI-R-BT1 is a subtracted library derived from a mixture of
the following tissues: hippocampus, thalamus, mid-brain,
medulla, corpus striatum, cerebral cortex and testis. For
a detailed description of the library from which this
clone was derived, please visit our web site at
rategen.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research

```

```

6:791-806, 1996)
TAG_LIB=UI-R-BT1
TAG_TISSUE=testis
TAG_SEQ=ACGCAG"
BASE COUNT 165 a 98 c 96 g 125 t
ORIGIN

Query Match 38.8%; Score 295.2; DB 9; Length 484;
Best Local Similarity 76.7%; Pred. No. 2.7e-32;
Matches 376; Conservative 0; Mismatches 108; Indels 6; Gaps 1;

QY 258 agatcctgtttgggaattatgaccttttctttggaggttatctcttcttctactttgttaa 317
Db 484 AGATCCTCTTTGGGAATCATGAACCTCTCATTTGGAGTGTGTTCCCTTTACCTTGGTAA 425
QY 318 aaccatatccaaagtttccctttatttcttcaggtatocattctgggctctgttt 377
Db 424 ACCCATACCCAAAGTTCCCTTTATATTTCTCAGGATATCCTTTCTGGGCTCTGCTT 365
QY 378 ttttcattatttgggagccttctcatttgcagtgaaagaaacacacagaactctga 437
Db 364 TGTTTCATTAACCTCTGGGACCTTTCTGATTGCTCTGAAAAAGAAAAACTACAGACACTCTGA 305
QY 438 taatattgagccgaataatgaattttctttagtgccttgagcaatagctggaatcattc 497
Db 304 TAAAAATGAGCCAGCGATGAATTTACTTAGTCCCTGGGAGCAGCAGCTGGGAATCATTC 245
QY 498 tcttcatttgggttctcattcatttgcagtaacaaactacatttgggttattctcaccacaata 557
Db 244 TCCTCATATGTGCTTCTCTTCTAGATGGGAATTCATCTGTGGCTATTCTCCAGATGGTA 185
QY 558 gtcagtgtaagctgttactgtctcttcttgggaatttctgattacattgacttcca 617
Db 184 TTCAGTGTGGTGCTATTACCACTCTATTGTTGGGATTTGATTATGTTGATGATCTCA 125
QY 618 gcattattgaattattcatttctccttctccttctcatttgggtgcacacagagatt 677
Db 124 CGTTCGTCGAACTGTTTCATTCCTCTCTTTTCCTCGATTTCCTGGGTCGCTACTCAGAGAAA 65
QY 678 gtgattgtgaacaattgttgcactgacactgagaaataaagatgtgttaaaatctcaa 737
Db 64 GTGGGGG-----ATGCTGTTGAATAGCACTGTACAAAGTAAACTGCAAAATTGAAAAAAA 11
QY 738 aaaaaa 747
Db 10 AAAAAAAA 1

RESULT 9
AI002083/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
ot38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
3', mRNA sequence.
AI002083
AI002083.1 GI:3202120
EST.
human.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 482)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 393.

FEATURES

Location/Qualifiers
1..482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH108"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
157 a 85 c 99 g 141 t
BASE COUNT
ORIGIN
Query Match 35.7%; Score 271; DB 9; Length 482;
Best Local Similarity 71.8%; Pred. No. 6.3e-29;
Matches 479; Conservative 0; Mismatches 0; Indels 188; Gaps 2;
QY 66 tcaatccttcaaatatcacgcacacatcatgattcagcagcagcacagtcg 125
|||||
DB 482 TCATCTCCTTCAATATACCGACCATCATGATTCAGCACCGCACAGTCGG 423
QY 126 tgttctggtattctccagaaatcaatgcttcagaatatgagtcacagaaattcag 185
|||||
DB 422 TGTTCTGTTATTCCTCCAGAAATCACTGCTTCAGATATGATGATCCAGAACTTCAG 363
QY 186 ccagacatttcaactcaaaccttgcaaaaattattgctagaaaaatgaaatct 245
DB 362 CCAGGACCTTTTCACTCAAGCCCTTCGAAAAATTTATTCGTAGAAAAATGAAATCT 303
QY 246 taggagcatccagatcctgttggaaattgacatttcttttggagttatcttcctt 305
|||||
DB 302 TAGS----- 297
QY 306 tcaattgtaaaaccatatccaaggttccctttatatttttttcagagatcattct 365
DB 298 -----GGATATCATTTCT 286
QY 366 ggggctctgttttcaattatctggagccttctcaattgagtgaaagaaaccca 425
DB 285 GGGCTCTGTTTGTTCATTAATCTGAGCCCTTCTTAATTCAGTGAAGAAACCA 226
QY 426 cagaaactctgataatattgagccgaataatgaatttcttctgcccctggagcaatag 485
|||||
DB 225 CAGAAACTCT----- 214
QY 486 ctggaatcattctcctcaatttggtttcattctcctagatcaaaactacatttggtatt 545
DB 215 -----GATCAAAACTACATTTGTGTTATT 191
QY 546 ctcacaaaaatagtcagtgtaagcgtgttactgtctcttcttgggaattttgattacat 605
DB 190 CTCACCAAAATAGTCAGTGAAGCGTGTACTGCTGTTCTTGGGAATTTGATTACAT 131
QY 606 tgaactttcaggaattgaattattcattctctgcttctcatttttgggggtgcc 665
DB 130 TGATGACTTTTTCAGCATTTATGAATTAATTCATTTCTGCTTTCTCAATTTTGGGTGCC 71
QY 666 actcagaggttgattggaacaattgttggactagcactgtgagaaatgagtg 725
|||||

Db 70 ACTCAGAGGATTGTGATTGTGAACAATGTTGTGACTAGCAGTGTGAGAAATAAGATGTG 11
QY 726 ttaaaat 732
|||||
Db 10 TAAAAAT 4
RESULT 10
AA411806/c 410 bp mRNA linear EST 12-AUG-1997
LOCUS z667a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION AA411806
VERSION AA411806.1 GI:2070377
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 410)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
, Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
, T., Waterston,R. and Wilson,K.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: -4ml3 fwd. ET from Amersham
High quality sequence stop: 401.
Location/Qualifiers
1..410
/organism="Homo sapiens"
/db_xref="GDB:5924284"
/db_xref="taxon:9606"
/clone="IMAGE:727372"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
147 a 68 c 76 g 119 t
BASE COUNT
ORIGIN
Query Match 35.2%; Score 267.4; DB 9; Length 410;
Best Local Similarity 96.1%; Pred. No. 2.2e-28;
Matches 274; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 178 acttcagccacgaccttttcaactcaagcccttcgcaaaattatttgcgtagaaaaat 237
|||||
Db 410 ACTTTCAGCCAGCAGCTTTTCAACTCAAGCCCTTGCAGAAATATTATTCGTAGAAAAAT 351
QY 238 gaaatctttaggagcatccagatcctctgttggaaattatgacctttcttttggagttat 297
|||||
Db 350 GAAAAATCTAGGGACTATCCAGATCCCTGTTTGGAAATTATGACCTTTTCTTTTGGAGTTAT 291
|||||

| | | | |
|-----------------------|-----|---|--|
| Qy | 260 | atcctgtttgggaattacatgacaccttttcttttggaggttatotctcttttactttgttaaaa | 319 |
| Db | 329 | ATCCTGTTTGGGAATATGACCTTTTCTTTTGGAGTTATCTTCCTTTTCACTTGTGTAATA | 270 |
| Qy | 320 | ccatacccaaggttcccttttatatttcttcaggatcacattctggggctctgttttg | 379 |
| Db | 269 | CGATATCCAGGTTTCCCTTTATATTCTTTCAGGATATCCATTCTGGGGCTCTGTGTTG | 210 |
| Qy | 380 | tccattaattctggagccttccctaatctgcagtgaaagaaaaaacccagagaaactctgata | 439 |
| Db | 209 | TTTCATTTAAATTCGTGGAGCCTCTCTAATTCAGTGAAGAAAAACCAACAGAAAACTCGGGA | 150 |
| Qy | 440 | atattgagccgaataatgaattt | 462 |
| Db | 149 | ATTTTGATTGATTCATGACTTT | 127 |
| RESULT | 12 | | |
| AQ108532 | | | |
| LOCUS | | 538 bp | DNA |
| DEFINITION | | CIT-HSP-2379G7.TF | CIT-HSP Homo sapiens genomic clone 2379G7, DNA |
| ACCESSION | | AQ108532 | |
| VERSION | | AQ108532.1 | GI:3485222 |
| KEYWORDS | | GSS. | |
| SOURCE | | human. | |
| ORGANISM | | Homo sapiens | |
| REFERENCE | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| AUTHORS | | 1 (bases 1 to 538) | |
| | | Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C. | |
| TITLE | | Use of a random human BAC End Sequence Database for Sequence-Ready Map Building | |
| JOURNAL | | Unpublished (1998) | |
| COMMENT | | Other_GSSs: CIT-HSP-2379G7.TR | |
| | | Contact: Mark Adams | |
| | | Department of Eukaryotic Genomics | |
| | | The Institute for Genomic Research | |
| | | 9712 Medical Center Dr., Rockville, MD 20850, USA | |
| | | Tel: 301 838 0200 | |
| | | Fax: 301 838 0208 | |
| | | Email: mdadams@tigr.org | |
| | | Clones are available from Research Genetics (info@resgen.com). BAC | |
| | | end search page: | |
| | | http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. | |
| | | Seq primer: M13-21 | |
| | | Class: BAC ends. | |
| FEATURES | | | |
| source | | 1..538 | |
| | | /organism="Homo sapiens" | |
| | | /db_xref="taxon:9606" | |
| | | /clone="2379G7" | |
| | | /clone_lib="CIT-HSP" | |
| | | /sex="Male" | |
| | | /cell_type="Sperm" | |
| | | /note="vector: pBelBAC11; Site_1: HindIII; Site_2: HindIII" | |
| BASE COUNT | | 140 a 108 c 115 g 175 t | |
| ORIGIN | | | |
| Query Match | | 19.5%; | Score 148.4; DB 12; Length 538; |
| Best Local Similarity | | 93.4%; | Pred. No. 6.8e-12; |
| Matches | | 155; Conservative | 0; Mismatches 11; Indels 0; Gaps 0; |
| Qy | 571 | tgtactgtcgtttctgggaattttgataactgtacattgatacttcagattatgaatt | 630 |
| Db | 152 | TCTCTCTCTATTACAGGGAATTTGATTACATTGATGACTTTCAGCATTTATGAAAT | 211 |
| Qy | 631 | attcatttctcgtcttctccaattttgggggtgcacctcagaggattgtgattggaaca | 690 |

```

Db 212 ATTCAATTCCTGCGCTTCTCAATTTGGGTGGCCACTCAGAGGATGTGATCTGAACA 271
QY 691 atgtgttgactagcactgtgagaataaagtgtgtaaaatctca 736
LOCUS |||||
Db 272 ATGTTGTGACTAGCAGCTGTGAGATAAAGATGTGTAATAATATTA 317
LOCUS |||||

RESULT 13
AA707529 281 bp mRNA linear EST 12-JAN-1999
LOCUS aa707529.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292062 3',
DEFINITION mRNA sequence.
ACCESSION AA707529
VERSION AA707529.1 GI:2717447
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 281)
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 456 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 218.
Location/Qualifiers
FEATURES
source
1. .281
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1292062"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - Oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 94 a 54 c 45 g 88 t
ORIGIN
Query Match 19.1%; Score 145; DB 9; Length 281;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 tgggaattttgattacattgacatttcagcattattgaattatttccttcgctt 647
Db 148 TGGGAATTTTATTACATTGATGACATTCAGCATTTATTAATTTCTCTGCGCTT 89
LOCUS |||||
QY 648 tctcaattttgggtgcactcagagattgtgattgtgaacaattgttgactagcac 707
LOCUS |||||
Db 88 TCTCAATTTTGGGTGGCCACTCAGAGGATGTGATTGTGAACAAGTTGTTGACTAGCAC 29
LOCUS |||||
QY 708 tctgagaataaagatgtgttaaaat 732
LOCUS |||||
Db 28 TGTGAATAAAGATGTGTTAAAT 4
LOCUS |||||

```

```

RESULT 14
BF319786/c
LOCUS BF319786
DEFINITION uy63h11.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
IMAGE:3664293 3', mRNA sequence.
ACCESSION BF319786
VERSION BF319786
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 332)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
CONTACT: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1425061
High quality sequence stop: 313.
Location/Qualifiers
FEATURES
source
1. .332
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
/clone_lib="McCarrey Eddy round spermatid"
/clone="IMAGE:3664293"
/sex="male"
/tissue_type="round spermatids, pooled from multiple mice"
/dev_stage="60 day"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pBluescript SK+ (Stratagene
); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dt-primed
[5'-(GA)10-AC TAGTCGAGTGTGTTTTTTTTTTT-3'] and directionally
cloned using 5' linkers 5'-AATTCGCGACGAG-3' and
5'-CTCGTGGCCG-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from lambda-UnizAP-XR) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 98.5% recombinants.
References: J. Androl. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D.
(Southwest Foundation for Biomedical Research, Dept. of
Genetics); excision done by E.M. Eddy, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences). Original lambda-based library is
available through ATCC, catalog #63423."
BASE COUNT 99 a 75 c 63 g 95 t
ORIGIN

```

```

Query Match 19.0%; Score 144.2; DB 10; Length 332;
Best Local Similarity 76.9%; Pred. No. 3.5e-11;
Matches 176; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 379 gtctcaattcttgagccttccttaattgcagtgaaagaaacacacgaactctgat 438
LOCUS |||||
Db 229 GTTCATTAACTCTGGGCGCTTCTGATTCTCTGAAAGAAAACACTACGACACTCIGAT 170
LOCUS |||||
QY 439 atattggccgcaataatgaattttttagtcgccctgggagcaatagctggaattctt 498
LOCUS |||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 11:02:16 ; Search time 26.58 Seconds
(without alignments)
1301.692 Million cell updates/sec

Title: US-09-821-821-2

Perfect score: 1027

Sequence: 1 MDSSSTAHPVFLVFPPEITA.....SLPFSILGCHSDCDEQCC 200

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1027 | 100.0 | 200 | 4 | Q9H3V2 |
| 2 | 1021 | 99.4 | 200 | 4 | Q9BZH1 |
| 3 | 202.5 | 19.7 | 205 | 4 | Q9H3V3 |
| 4 | 202.5 | 19.7 | 220 | 4 | Q9BY18 |
| 5 | 202.5 | 19.7 | 239 | 4 | Q96JQ5 |
| 6 | 200.5 | 19.5 | 197 | 4 | Q9P1S3 |
| 7 | 188.5 | 18.4 | 217 | 11 | Q9D7Z9 |
| 8 | 187.5 | 18.3 | 217 | 11 | Q9N08 |
| 9 | 185.5 | 18.1 | 214 | 4 | Q96HJ5 |
| 10 | 183.5 | 17.9 | 225 | 4 | Q96PG6 |
| 11 | 183.5 | 17.9 | 248 | 4 | Q9H2W1 |
| 12 | 182.5 | 17.8 | 225 | 4 | Q9HC76 |
| 13 | 180 | 17.5 | 226 | 11 | Q9ES61 |
| 14 | 178.5 | 17.4 | 225 | 11 | Q9N05 |
| 15 | 176 | 17.1 | 268 | 11 | Q9N10 |
| 16 | 174 | 16.9 | 226 | 11 | Q9D3F6 |

| | | | | | | |
|----|-------|------|-----|----|--------|--------------------|
| 17 | 172.5 | 16.8 | 220 | 11 | Q9D8E5 | Q9d8e5 mus musculu |
| 18 | 172.5 | 16.8 | 267 | 4 | Q9NXJ0 | Q9nxj0 homo sapien |
| 19 | 171 | 16.7 | 178 | 4 | Q9H3V1 | Q9h3v1 homo sapien |
| 20 | 171 | 16.7 | 679 | 4 | Q96JA4 | Q96ja4 homo sapien |
| 21 | 168 | 16.4 | 249 | 4 | Q9H2N3 | Q9h2n3 homo sapien |
| 22 | 167 | 16.3 | 268 | 11 | Q9EQZ0 | Q9eqz0 mus musculu |
| 23 | 165 | 16.1 | 135 | 6 | Q9SLT4 | Q9slt4 macaca fasc |
| 24 | 160.5 | 15.6 | 213 | 11 | Q920C4 | Q920c4 mus musculu |
| 25 | 160.5 | 15.6 | 240 | 4 | Q9GZW8 | Q9gzw8 homo sapien |
| 26 | 158.5 | 15.4 | 247 | 11 | Q9N07 | Q9n07 mus musculu |
| 27 | 156 | 15.2 | 244 | 11 | Q9N09 | Q9n09 mus musculu |
| 28 | 155.5 | 15.1 | 247 | 11 | Q9D0X1 | Q9d0x1 mus musculu |
| 29 | 154 | 15.0 | 244 | 11 | Q9D8W9 | Q9d8w9 mus musculu |
| 30 | 148 | 14.4 | 249 | 4 | Q9H2L1 | Q9h2l1 homo sapien |
| 31 | 145.5 | 14.2 | 176 | 11 | Q9DC76 | Q9dc76 mus musculu |
| 32 | 145.5 | 14.2 | 250 | 4 | Q9BY19 | Q9by19 homo sapien |
| 33 | 139.5 | 13.6 | 190 | 11 | Q9EQV7 | Q9eqv7 mus musculu |
| 34 | 137.5 | 13.4 | 124 | 11 | Q99MX6 | Q99mx6 mus musculu |
| 35 | 137.5 | 13.4 | 167 | 4 | Q96PG7 | Q96pg7 homo sapien |
| 36 | 137.5 | 13.4 | 230 | 11 | Q9EQY9 | Q9eqy9 mus musculu |
| 37 | 136 | 13.2 | 234 | 11 | Q99N04 | Q99n04 mus musculu |
| 38 | 133 | 12.0 | 125 | 11 | Q912V5 | Q91zy5 mus musculu |
| 39 | 108.5 | 10.6 | 215 | 11 | Q9D2W6 | Q9d2w6 mus musculu |
| 40 | 106 | 10.3 | 307 | 16 | Q9CES9 | Q9ces9 lactococcus |
| 41 | 103 | 10.0 | 175 | 11 | Q9D7R4 | Q9d7r4 mus musculu |
| 42 | 103 | 10.0 | 455 | 5 | Q19442 | Q19442 caenorhabdi |
| 43 | 100 | 9.7 | 152 | 4 | Q14298 | Q14298 homo sapien |
| 44 | 96 | 9.3 | 387 | 16 | Q9PNQ1 | Q9pnq1 campylobact |
| 45 | 95.5 | 9.3 | 355 | 17 | Q96ZJ3 | Q96zj3 sulfolobus |

ALIGNMENTS

RESULT 1

Q9H3V2 ID Q9H3V2 PRELIMINARY; PRT; 200 AA.

AC Q9H3V2;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE MS4A5 (MS4A5 PROTEIN).

GN MS4A5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Ishibashi K., Sasaki S., Marumo F.;

RT "Cloning of three CD20 homolog from human, putative calcium

channels";

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21295030; PubMed=11401424;

RA Liang Y., Tedder T.F.;

RT "Identification of a CD20-, FcepsilonRIbeta-, and HTM4-Related Gene

Family: Sixteen New MS4A Family Members Expressed in Human and

Mouse.";

RL Genomics 72:119-127(2001).

DR EMBL; AB013103; BAB18739.1; -

DR EMBL; AF237907; AAK37416.1; -

SQ SEQUENCE 200 AA; 22283 MW; F9282E5D15BC5514 CRC64;

Query Match 100.0%; Score 1027; DB 4; Length 200;

Best Local Similarity 100.0%; Pred. No. 1.2e-84;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSSTAHPVFLVFPPEITA.....SLPFSILGCHSDCDEQCC 200

|||||

| | | | |
|--|---|---|--------------|
| Db | 1 | MDSSTAHPVFLVFPPEITASEYESTELSATTFTSQSPLOKLFARKMKILGTIQLFGIM | 60 |
| QY | 61 | TFSGVIFLTLKPYPRPFIFLSGYFPGSVLFINSAGFLIAVKRKTETLILSRIM | 120 |
| Db | 61 | TFSGVIFLTLKPYPRPFIFLSGYFPGSVLFINSAGFLIAVKRKTETLILSRIM | 120 |
| QY | 121 | NFLSALGAAGIILLTFTGFILDQNYICGYSHQNSQCKAVTVFLGILITLMTFSIIELFI | 180 |
| Db | 121 | NFLSALGAAGIILLTFTGFILDQNYICGYSHQNSQCKAVTVFLGILITLMTFSIIELFI | 180 |
| QY | 181 | SLPFSILGCHSEDCDCQCC 200 | |
| Db | 181 | SLPFSILGCHSEDCDCQCC 200 | |
| RESULT | 2 | | |
| Q9BZH1 | | PRELIMINARY; | PRT; 200 AA. |
| AC | Q9BZH1 | | |
| DT | 01-JUN-2001 | (TEMBLrel. 17, Created) | |
| DT | 01-JUN-2001 | (TEMBLrel. 17, Last sequence update) | |
| DT | 01-JUN-2001 | (TEMBLrel. 17, Last annotation update) | |
| DE | TESTIS-EXPRESSED TRANSMEMBRANE-4 PROTEIN. | | |
| GN | TETM4. | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| OX | NCBI_TaxID=9606; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=21092614; PubMed=11162526; | | |
| RA | Hulett M.D., Pagler E., Hornby J.R., Hogarth P.M., Eyre H.J., | | |
| RA | Baker E., Crawford J., Sutherland G.R., Ohms S.J., Parish C.R.; | | |
| RT | "Isolation, tissue distribution, and chromosomal localization of a | | |
| RT | novel testis-specific human four-transmembrane gene related to CD20 | | |
| RT | and Fcpsi1onRI-beta."; | | |
| RL | Biochem. Biophys. Res. Commun. 280:374-379(2001). | | |
| DR | EMBL; AF321127; AAK01641.1; -. | | |
| KW | Transmembrane. | | |
| SQ | SEQUENCE 200 AA; 22249 MW; 7B282E5D15B25BF4 CRC64; | | |
| Query Match 99.4%; Score 1021; DB 4; Length 200; | | | |
| Best Local Similarity 99.5%; Pred. No. 4.3e-84; | | | |
| Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | | |
| QY | 1 | MDSSTAHPVFLVFPPEITASEYESTELSATTFTSQSPLOKLFARKMKILGTIQLFGIM | 60 |
| Db | 1 | MDSSTAHPVFLVFPPEITASEYESTELSATTFTSQSPLOKLFARKMKILGTIQLFGIM | 60 |
| QY | 61 | TFSGVIFLTLKPYPRPFIFLSGYFPGSVLFINSAGFLIAVKRKTETLILSRIM | 120 |
| Db | 61 | TFSGVIFLTLKPYPRPFIFLSGYFPGSVLFINSAGFLIAVKRKTETLILSRIM | 120 |
| QY | 121 | NFLSALGAAGIILLTFTGFILDQNYICGYSHQNSQCKAVTVFLGILITLMTFSIIELFI | 180 |
| Db | 121 | NFLSALGAAGIILLTFTGFILDQNYICGYSHQNSQCKAVTVLLGLLITLMTFSIIELFI | 180 |
| QY | 181 | SLPFSILGCHSEDCDCQCC 200 | |
| Db | 181 | SLPFSILGCHSEDCDCQCC 200 | |
| RESULT | 3 | | |
| Q9H3V3 | | PRELIMINARY; | PRT; 205 AA. |
| ID | Q9H3V3 | | |
| AC | Q9H3V3 | | |
| DT | 01-MAR-2001 | (TEMBLrel. 16, Created) | |
| DT | 01-MAR-2001 | (TEMBLrel. 16, Last sequence update) | |
| DT | 01-DEC-2001 | (TEMBLrel. 19, Last annotation update) | |
| DE | MS4A4. | | |
| GN | MS4A4. | | |
| OS | Homo sapiens (Human). | | |

| | | |
|--|---|---------------------------|
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=21142397; PubMed=11245982; | |
| RA | Ishibashi K., Suzuki M., Sasaki S., Imai M.; | |
| RT | "Identification of a new multigene four-transmembrane family (MS4A) | |
| RT | related to CD20, Htm4 and beta subunit of the high-affinity IgE | |
| RT | receptor."; | |
| RL | Gene 264:87-93(2001). | |
| DR | EMBL; AB013102; BAB18738.1; -. | |
| SQ | SEQUENCE 205 AA; 21790 MW; FDB12899FCC551B3 CRC64; | |
| Query Match 19.7%; Score 202.5; DB 4; Length 205; | | |
| Best Local Similarity 31.8%; Pred. No. 1.1e-10; | | |
| Matches 49; Conservative 31; Mismatches 71; Indels 3; Gaps | | |
| QY | 39 LQKLFAR-KMKILGTIQLFGIMTFSFGVIFLTLKPYPRPFIFLSGYFPGSVLFIN 97 | |
| Db | 20 LQKFLKGPVKLVGVVQILTALMSLSMGITMCMASNTYGSNPISVIGYITWGSVMFII 79 | |
| QY | 98 SGAFLIAVKRKTETLILSRIMNFLSALGAAGIILLTFF--GFILDQNYICGYSHQNSQ 155 | |
| Db | 80 SGSLSTAAGIRTKGLVRGSLGNMTSSVLAASGILINTFSLAFYFHHPCNYCYGNSNN 139 | |
| QY | 156 CKAVTVFLGILITLMTFSIIELFISLPFSILGC 189 | |
| Db | 140 CHGTMSILMGDLGMYLLSVLEFCIAVSLSAFGC 173 | |
| RESULT 4 | | |
| Q9BY18 | | |
| ID | Q9BY18 | PRELIMINARY; PRT; 220 AA. |
| AC | Q9BY18; | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Created) | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | |
| DE | MS4A4A PROTEIN. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RC | TISSUE=PLACENTA; | |
| RX | MEDLINE=21295030; PubMed=11401424; | |
| RA | Liang Y., Tedder T.F.; | |
| RT | "Identification of a cd20-, fcepslonribeta-, and htm4-related gene | |
| RT | family; sixteen new ms4a family members expressed in human and | |
| RT | mouse."; | |
| RL | Genomics 72:119-127(2001). | |
| DR | EMBL; AF237912; AAK37594.1; -. | |
| SQ | SEQUENCE 220 AA; 23220 MW; 7648C762197C1179 CRC64; | |
| Query Match 19.7%; Score 202.5; DB 4; Length 220; | | |
| Best Local Similarity 31.8%; Pred. No. 1.2e-10; | | |
| Matches 49; Conservative 31; Mismatches 71; Indels 3; Gaps | | |
| QY | 39 LQKLFAR-KMKILGTIQLFGIMTFSFGVIFLTLKPYPRPFIFLSGYFPGSVLFIN 97 | |
| Db | 35 LQKFLKGPVKLVGVVQILTALMSLSMGITMCMASNTYGSNPISVIGYITWGSVMFII 94 | |
| QY | 98 SGAFLIAVKRKTETLILSRIMNFLSALGAAGIILLTFF--GFILDQNYICGYSHQNSQ 155 | |
| Db | 95 SGSLSTAAGIRTKGLVRGSLGNMTSSVLAASGILINTFSLAFYFHHPCNYCYGNSNN 154 | |
| QY | 156 CKAVTVFLGILITLMTFSIIELFISLPFSILGC 189 | |
| Db | 155 CHGTMSILMGDLGMYLLSVLEFCIAVSLSAFGC 188 | |

RESULT
Q99N08

```

ID Q99N08 PRELIMINARY; PRT; 217 AA.
AC Q99N08;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MSA6C PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TOTAL FETUS;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilonribeta-, and htm4-related gene
family: sixteen new ms4a family members expressed in human and
mouse."
RL Genomics 72:119-127(2001).
DR EMBL; AF237910; AAK37419.1; -.
SQ SEQUENCE 217 AA; 23622 MW; 8BF521AF22DBB7BD CRC64;

Query Match 18.3%; Score 187.5; DB 11; Length 217;
Best Local Similarity 26.0%; Pred. No. 2.5e-09;
Matches 54; Conservative 41; Mismatches 80; Indels 33; Gaps 5;

QY 2 DSSTAHPVFLVFPPEITASEYESTELSATFTSQPLQKLFARKMKILGTIQLFGIMT 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 ETITISPNGINFPQK-----DESQPTQQRQSLKHLKAEIKVIAIQIMCAVTV 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 FSGVIFLTLKPYRPFPIF---LSGYPFWGSLVFLNSGAFLIAVKKTKTETILII 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 LALGII--LASVPPVPYFNSVLLKSGYPFGIFAGLFFIASGILSIITKSKPLVDAS 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 118 RIMNFLSALGAIGIILLTGF-----ILDQNYICG-YSHQNSOCKAV 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 LTNILSVSAFVGIIISVSLAGLHPASEQCKQSKLSLIEHDHYQPFYNSRSECAVT 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 160 TVFLGILITLMTFSIIEFLSILPFSIL 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 KSLITGALSVMILISVLELGLALLSAML 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
Q96HJ5 PRELIMINARY; PRT; 214 AA.
AC Q96HJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:14809).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008487; AAH08487.1; -.
SQ SEQUENCE 214 AA; 22933 MW; D30C276213DB8F3D CRC64;

Query Match 18.1%; Score 185.5; DB 4; Length 214;
Best Local Similarity 28.7%; Pred. No. 3.8e-09;
Matches 60; Conservative 39; Mismatches 85; Indels 25; Gaps 8;

QY 1 MDSSTAHPVFLVFPPEITASEYESTELSATFTSQPLQKLFARKMKILGTIQLFGI 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 LGSASAHGTPGSEAGPEELNTSVYQPDIGS-----PDYQK---AKLQVLGAIQILNAA 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 60 MTFSGVIFLFTLLKPY-----PRFPFIFLSGYPFWGSLVFLNSGAFLIAVKKTKTETILII 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 MILALGV-FLGSLQYPYHQKHFHFFFTVGTPIWGAFFCGSSGLSVVAGIKPRTWIQ 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 116 LSRIMNFLSALGAIGIILLTGFILDQNYI--CCYSHQNSQ----CKAVTVFLGILIT 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 NSFGMNIASATATLVGTAFSLNIAVNIQSLRSC---HSSSESDLCNMGISNGMVSL 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 170 LMTFSIIEFLSILPFSILGCHSEDCDCEQ 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 LLILLLLELCVTISTIAMWCNANCCNSRE 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q96PG6 PRELIMINARY; PRT; 225 AA.
AC Q96PG6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MSA4A6A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21382044; PubMed=11486273;
RA Liang Y., Buckley T.R., Tu L., Langdon S.D., Tedder T.F.;
RT "Structural organization of the human MS4A gene cluster on Chromosome
11q12."
RL Immunogenetics 53:357-368(2001).
DR EMBL; AF354930; AAL07357.1; -.
SQ SEQUENCE 225 AA; 24331 MW; 12FBF5DCDC71B56A CRC64;

Query Match 17.9%; Score 183.5; DB 4; Length 225;
Best Local Similarity 26.4%; Pred. No. 6e-09;
Matches 55; Conservative 39; Mismatches 91; Indels 23; Gaps 6;

QY 1 MDSSTAHPVFLVFPPEITASEYESTELSATFTSQPLQKLFARKMKILGTIQLFGIM 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MTSQVPNETIIVLPSNVI--NFSQAERPEPTNOGQDSLKKHLHAEIKVIGTIQLCGMM 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 TFSGVIFLFTLLKP-YPRFPFIFL-SCYPEWGSVFLNSGAFLIAVKKTKTETILILSR 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 VLSLGIILASASFSNFTQVTVSTLLNSAYPIGFFFFIISGLSIATKRLKLVHSSL 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 119 IMNFLSALGAIGIILL-----TGFILDQN-----YICGYSHQN---SQCKAV 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 VGSILSALSALVGFIILSVKQATLNPASLQCELDKNIPTRSYVSFYVHDSLYITDCVTA 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 160 TVFLGILITLMTFSIIEFLSILPFSIL 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 KASLAGTILSLICTLLEFLCLAVITAVL 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q9H2W1 PRELIMINARY; PRT; 248 AA.
AC Q9H2W1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDA01 (MSA4A6A-POLYMORPH) (MS4A6A PROTEIN).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PHEOCHROMOCYTOMA;

```

```
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.,
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP .SEQUENCE FROM N.A.
RC TISSUE=LIVER, AND SPLEEN;
RX MEDLINE=21295030; PubMed=11401424;
RLiang Y., Tedder T.F.;
RT "Identification of a CD20-, FcepsilonRIbeta-, and HTm4-Related Gene
RT Family: Sixteen New MS4A Family Members Expressed in Human and
RT Mouse.;"
RL Genomics 72:119-127(2001).
DR EMBL; AF212240; AAG41780.1; -.
DR EMBL; AF286866; AAK37994.1; -.
DR EMBL; AF237908; AAK37417.1; -.
SQ SEQUENCE 248 AA; 26943 MW; 49892BC25D220A23 CRC64;

Query Match 17.9%; Score 183.5; DB 4; Length 248;
Best Local Similarity 26.4%; Pred. No. 6.6e-09;
Matches 55; Conservative 39; Mismatches 91; Indels 23; Gaps 6;

Qy 1 MDSSTAHPVLPVFPPEITASEYSTELSATTFSTQSPLOKLFARKMKILGTIQLFGIM 60
Db 1 MTSQVPNETIIVLPSNVI--NFSQAEKPEPTNQGDLSKKHLHAEIKVIGTIQLCGMM 58
Qy 61 TFSFGVIFLTLKP-VPRPFPIFL-SGYFPWGSVLFNSGAFLIAVKKRTTETLIILSR 118
Db 59 VLSGLIILASFSFNFTQVITLNSAYPIGPFPIISGSLSIATEKRLKLVHSSL 118
Qy 119 IMNFSALGAIAGIILL-----TFGFILDQN-----YICGYSHQN---SQCKAV 159
Db 119 VGSILSALSALVGFIILSVKQATLNPSALQCELDKNNIPTRSYVSYFYHDSLYTDCYTA 178
Qy 160 TVLFGILITLMTFSIIEFLISLPSIL 187
Db 179 KASLAGTSLMLICTLLEFLCLAVLTAFL 206

RESULT 12
Q9HC76 PRELIMINARY; PRT; 225 AA.
AC Q9HC76;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE CD20-LIKE PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang W., Wan T., He L., Yuan Z., Cao X.;
RL "Novel human CD20-like molecule.;"
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF142409; AAG27920.1; -.
SQ SEQUENCE 225 AA; 24317 MW; 144D2EB7DC71B56A CRC64;

Query Match 17.8%; Score 182.5; DB 4; Length 225;
Best Local Similarity 26.4%; Pred. No. 7.4e-09;
Matches 55; Conservative 39; Mismatches 91; Indels 23; Gaps 6;

Qy 1 MDSSTAHPVLPVFPPEITASEYSTELSATTFSTQSPLOKLFARKMKILGTIQLFGIM 60
Db 1 MTSQVPNETIIVLPSNVI--NFSQAEKPEPTNQGDLSKKHLHAEIKVIGTIQLCGMM 58
Qy 61 TFSFGVIFLTLKP-VPRPFPIFL-SGYFPWGSVLFNSGAFLIAVKKRTTETLIILSR 118
Db 59 VLSGLIILASFSFNFTQVITLNSAYPIGPFPIISGSLSIATEKRLKLVHSSL 118
Qy 119 IMNFSALGAIAGIILL-----TFGFILDQN-----YICGYSHQN---SQCKAV 159
```

```
Db 119 VGSILSALSALVGFIILSVKQATLNPSALQCELDKNNIPTRSYVSYFYHDSLYTDCYTA 178
Qy 160 TVLFGILITLMTFSIIEFLISLPSIL 187
Db 179 KASLAGTSLMLICTLLEFLCLAVLTAFL 206

RESULT 13
Q9ES61 PRELIMINARY; PRT; 226 AA.
AC Q9ES61;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHANDRA PROTEIN (MS4A4B PROTEIN).
GN LY116.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Venkataraman C., Schaefer G., Schindler U.;
RT "Chandra, a novel four transmembrane protein differentially expressed
RT in helper type I lymphocytes.;"
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=21295030; PubMed=11401424;
RLiang Y., Tedder T.F.;
RT "Identification of a CD20-, FcepsilonRIbeta-, and HTm4-Related Gene
RT Family: Sixteen New MS4A Family Members Expressed in Human and
RT Mouse.;"
RL Genomics 72:119-127(2001).
DR EMBL; AF230640; AAG09739.1; -.
DR EMBL; AF237913; AAK37595.1; -.
DR MGD; MGI:1913083; LY116.
SQ SEQUENCE 226 AA; 24090 MW; 78C6ED3E830AA116 CRC64;

Query Match 17.5%; Score 180; DB 11; Length 226;
Best Local Similarity 26.1%; Pred. No. 1.2e-08;
Matches 43; Conservative 36; Mismatches 62; Indels 24; Gaps 4;

Qy 40 QKLFARKMKILGTIQLFGIMTFSFGVIFLTLTKPKYPFPFFLGSYPFGWSVLFINS 99
Db 34 EKFLKGEPRVLGVQVIAINUSLGIILITLFLSELPTSVMLV---PWGSIIMFIVSG 90
Qy 100 AFLIAVKKRTTETLIIL-----LSRIMNFSALGAIAGIILLITFG--FILQNYICGYSHQN 153
Db 91 SLSIAAGVTPTKCLIVASLTNITSVLAATASIMGVSVAVGSQFPFRNY-----142
Qy 154 SQCKAVTVLFGILITLMTFSIIEFLISLPSILSGHSEDCDCEQ 198
Db 143 -----TITKGLDVLMLIFNMLEFLAVSFAFGCEASCCNSRE 180

RESULT 14
Q99N05 PRELIMINARY; PRT; 225 AA.
AC Q99N05;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MS4A4D PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

Search completed: August 27, 2002, 11:05:06
Job time: 170 sec

```
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=211295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilonribeta-, and htm4-related gene
RT family: sixteen new ms4a family members expressed in human and
RT mouse.";
RL Genomics 72:119-127(2001).
DR EMBL; AF237915; AAK37598.1; -.
SQ SEQUENCE 225 AA; 23859 MW; 9D3B2678787039B4 CRC64;

Query Match 17.4%; Score 178.5; DB 11; Length 225;
Best Local Similarity 31.4%; Pred. No. 1.7e-08;
Matches 50; Conservative 35; Mismatches 61; Indels 13; Gaps 6;

QY 40 OKLFARKKKILGTTQILGIMTFSEFGVIFLTLKPYPRPFIFLSGYPPWGSVLFINS 99
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 100 AFLIAVKKKTTETLIILSRIMNLSALGAIAGIILLTFGLDQNYICGYSHQNSOCKAV 159
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 TVLFGILITLMTFSIIFLFSILPFSILGCHSEDCDCEQ 198
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 145 A--SLDVLMTIL--NMLEFCIAVSVAFGCKACSCNSSE 179

RESULT 15
Q99N10 PRELIMINARY; PRT; 268 AA.
AC Q99N10;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MS4A8B PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=COLON;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilonribeta-, and htm4-related gene
RT family: sixteen new ms4a family members expressed in human and
RT mouse.";
RL Genomics 72:119-127(2001).
DR EMBL; AF237906; AAK37415.1; -.
SQ SEQUENCE 268 AA; 28568 MW; 490C79BF3BA1579C CRC64;

Query Match 17.1%; Score 176; DB 11; Length 268;
Best Local Similarity 27.1%; Pred. No. 3.3e-08;
Matches 60; Conservative 36; Mismatches 91; Indels 34; Gaps 9;

QY 1 MDSSTHSP----VFLVFPPEI-----TASE---YESTELSATTF-----TQS 37
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 MNTVTSPPGMANSVYVAPPNSYPVPGTVPMPIYPSNQPVHVISGHLPGIVPANTEP 76
QY 38 PLQKLFARKKKILGTTQILGIMTFSEFGVIFLTLKPYPRPFIFLSGYPPWGSVLFIN 97
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 PAQRVL--KRGQVILGAIQILGLVHIGLSIMTNLFSHYT--PVSLYGGFPFWGGIWFII 133
QY 98 SGAFLIAVKKKTTETLIILSRIMNLSALGAIAGIILLTFGLDQNYICGYSHQNSOCKAV 153
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 SGSLVAAEFQPNPCLNGSVGLNIFSAICSAVGIML-----FITDISSGTIYPSYTP 189
QY 154 SQKAVTVLFLGILITLMTFSIIFLFSILPFSILGCHSEDC 194
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 YQENLGVRTGVAISSVLLIFCLLELSIASVSSHFGQVACC 230
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 11:02:16 ; Search time 27.48 Seconds
(Without alignments)
177.770 Million cell updates/sec

Title: US-09-821-821-2
Perfect score: 1027
Sequence: 1 MDSSTAHSPVLVFPPEITA.....SLPFSILGCHSDCCEQCC 200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 202.5 | 19.7 | 239 | 2 | US-08-916-902A-1 |
| 2 | 202.5 | 19.7 | 239 | 2 | US-09-213-389-1 |
| 3 | 181.5 | 17.7 | 214 | 1 | US-08-318-492-4 |
| 4 | 181.5 | 17.7 | 214 | 1 | US-08-707-340-4 |
| 5 | 181.5 | 17.7 | 214 | 2 | US-08-916-902A-3 |
| 6 | 181.5 | 17.7 | 214 | 2 | US-08-994-578-4 |
| 7 | 181.5 | 17.7 | 214 | 2 | US-09-213-389-3 |
| 8 | 158 | 15.4 | 243 | 1 | US-07-869-933-29 |
| 9 | 158 | 15.4 | 243 | 1 | US-07-869-933-33 |
| 10 | 158 | 15.4 | 243 | 1 | US-08-201-879A-4 |
| 11 | 158 | 15.4 | 243 | 2 | US-08-916-902A-4 |
| 12 | 158 | 15.4 | 243 | 2 | US-09-213-389-4 |
| 13 | 158 | 15.4 | 243 | 4 | US-09-103-663-29 |
| 14 | 158 | 15.4 | 243 | 4 | US-09-103-663-33 |
| 15 | 158 | 15.4 | 246 | 1 | US-07-869-933-23 |
| 16 | 158 | 15.4 | 246 | 4 | US-09-103-663-23 |
| 17 | 157.5 | 15.3 | 235 | 1 | US-07-869-933-34 |
| 18 | 157.5 | 15.3 | 235 | 1 | US-08-201-879A-5 |
| 19 | 157.5 | 15.3 | 235 | 4 | US-09-103-663-34 |
| 20 | 123.5 | 12.0 | 244 | 1 | US-07-869-933-32 |
| 21 | 123.5 | 12.0 | 244 | 1 | US-08-201-879A-3 |
| 22 | 123.5 | 12.0 | 244 | 4 | US-09-103-663-32 |
| 23 | 87.5 | 8.5 | 956 | 2 | US-08-897-443-3 |
| 24 | 82 | 8.0 | 327 | 4 | US-08-748-506-24 |
| 25 | 79.5 | 7.8 | 371 | 2 | US-08-928-692-20 |
| 26 | 79.5 | 7.7 | 241 | 3 | US-08-808-148-1 |
| 27 | 79.5 | 7.7 | 241 | 4 | US-09-020-956-114 |

| | | | | | | |
|----|------|-----|------|---|-------------------|-------------------|
| 28 | 79.5 | 7.7 | 241 | 4 | US-09-030-607-114 | Sequence 114, App |
| 29 | 79.5 | 7.7 | 241 | 4 | US-09-439-313-114 | Sequence 114, App |
| 30 | 78.5 | 7.6 | 235 | 4 | US-09-247-155-126 | Sequence 126, App |
| 31 | 78.5 | 7.6 | 761 | 4 | US-09-235-451-4 | Sequence 4, Appli |
| 32 | 76.5 | 7.4 | 638 | 2 | US-08-897-443-1 | Sequence 1, Appli |
| 33 | 75.5 | 7.4 | 423 | 2 | US-08-494-907-14 | Sequence 14, Appl |
| 34 | 75.5 | 7.4 | 423 | 5 | PCT-US96-10986-14 | Sequence 14, Appl |
| 35 | 75 | 7.3 | 347 | 4 | US-09-087-889-14 | Sequence 14, Appl |
| 36 | 74.5 | 7.3 | 215 | 4 | US-09-087-232A-17 | Sequence 17, Appl |
| 37 | 74.5 | 7.3 | 352 | 4 | US-09-087-232A-13 | Sequence 13, Appl |
| 38 | 74.5 | 7.3 | 352 | 4 | US-08-861-105-14 | Sequence 14, Appl |
| 39 | 74.5 | 7.3 | 352 | 4 | US-08-575-967A-2 | Sequence 2, Appli |
| 40 | 74 | 7.2 | 1416 | 1 | US-08-061-465-4 | Sequence 4, Appli |
| 41 | 73.5 | 7.2 | 352 | 4 | US-09-045-583-52 | Sequence 52, Appl |
| 42 | 73.5 | 7.2 | 381 | 1 | US-08-467-125-2 | Sequence 2, Appli |
| 43 | 73.5 | 7.2 | 381 | 2 | US-08-911-320A-2 | Sequence 2, Appli |
| 44 | 73.5 | 7.2 | 381 | 4 | US-09-217-101-2 | Sequence 2, Appli |
| 45 | 73 | 7.1 | 360 | 4 | US-08-875-573-20 | Sequence 20, Appl |

ALIGNMENTS

RESULT 1
US-08-916-902A-1
; Sequence 1, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-1

| | | | | | | | | | |
|-----------------------|--------|--------------|----------|------------|--------|--------|----|------|--|
| Query Match | 19.7% | Score | 202.5; | DB 2; | Length | 239; | | | |
| Best Local Similarity | 31.8%; | Pred. No. | 2.4e-15; | | | | | | |
| Matches | 49; | Conservative | 31; | Mismatches | 71; | Indels | 3; | Gaps | |

| | | | |
|----|-----|---|-----|
| QY | 39 | LKLFAR-KMILGFIQLGFMTFSGVIEFTLLKLPPRPFFELSLGGVPFGSVLEFN | 97 |
| | | : : : : : : : : : : : : : | |
| Db | 54 | LOEKFLKEPKVLGVQQLTALMSLGSGLITMMCMASNYGNSPIISVIYGTTINGSWVFII | 113 |
| QY | 98 | SGAFLIAYKRRTETTLILSRINFLSALGAIGAILLTFF--GTILDONTICYGYSHQNS | 155 |
| | | : : : : : : : : : : : | |
| Db | 114 | SGSLSIAGAIRTTKGLVRGSLGNHTSSVLAASGILINTFSLAPSPHPHYCNYGNSNN | 173 |
| QY | 156 | CKANTVFLGLITYIMTFSTIEFLSPLPSILGC | 189 |
| | | : : : : : : : : : | |
| Db | 174 | CHGTSMILGDGMVLLISLVDFCIATVSIAFSAGCC | 207 |

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 19.7%; Score 202.5; DB 2; Length 239;
Best Local Similarity 31.8%; Pred. No. 2.4e-15;
Matches 49; Conservative 31; Mismatches 71; Indels 3; Gaps 2;

QY 39 LQKLFR-KMKILGTQLFGIMTFSEGVIFLFTLLKPYPRFFFLSGYPFWGSLFIN 97
||| : : : ||| : ||| : : : : : | : ||| ||| |||

| | | | |
|----|-----|--|-----|
| Db | 54 | LOEKFLKEBPKVLGVVVQILLFALMSLMSGMTMCMSASNTVGSNPTSVYIGTYINGSVYFII | 113 |
| QY | 98 | SGAFLIAYKRKTEFLIILSLXIMNFI.SALGAIAGIILLTF--GFTLOQNYICGYSHONSQ | 155 |
| Db | 114 | SGSLTAAGITRTKGLVRGSLGMMNTSSVLAAGILINTFLSLAFYSFHHPYCNYGNSNN | 173 |
| QY | 156 | KRAVTVLFGILITLMTFTSIELFISLPTFSILGC | 189 |
| Db | 174 | CHGTASILMGDLGMVLLLSVLEFCIAVSLSAFGC | 207 |

```

1  RESULT 3
2  US-08-318-492-4
3  / Sequence 4, Application US/08318492
4  / Patent No. 5552312
5  / GENERAL INFORMATION:
6  / APPLICANT: Lim, Bing
7  / APPLICANT: Adra, Chaker N.
8  / APPLICANT: Lelias, Jean-Michel
9  / TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
10 / TITLE OF INVENTION: ASSAYS
11 / NUMBER OF SEQUENCES: 4
12 / CORRESPONDENCE ADDRESS:
13 / ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
14 / STREET: Two Militia Drive
15 / CITY: Lexington
16 / STATE: Massachusetts
17 / COUNTRY: U.S.A.
18 / ZIP: 02173
19 / COMPUTER READABLE FORM:
20 / MEDIUM TYPE: Floppy disk
21 / COMPUTER: IBM PC compatible
22 / OPERATING SYSTEM: PC-DOS/MS-DOS
23 / SOFTWARE: PatentIn Release #1.0, Version #1.25
24 / CURRENT APPLICATION DATA:
25 / APPLICATION NUMBER: US/08/318,492
26 / FILING DATE:
27 / CLASSIFICATION: 530
28 / ATTORNEY/AGENT INFORMATION:
29 / NAME: Brook, David E.
30 / REGISTRATION NUMBER: 22,592
31 / REFERENCE/DOCKET NUMBER: BIH94-03
32 / TELECOMMUNICATION INFORMATION:
33 / TELEPHONE: (617) 861-6240
34 / TELEFAX: (617) 861-9540
35 / INFORMATION FOR SEQ ID NO: 4:
36 / SEQUENCE CHARACTERISTICS:
37 / LENGTH: 214 amino acids
38 / TYPE: amino acid
39 / TOPOLOGY: linear
40 / MOLECULE TYPE: protein
41 / IS-08-318-492-4

```

```

Query Match      17.7% ; Score 181.5 ; DB 1 ; Length 214 ;
Best Local Similarity 28.2% ; Pred. No. 5.2e-13 ;
Matches 59 ; Conservative 36 ; Mismatches 87 ; Indels 25 ; Gaps 8 ;

QY 1 MDSTAH-SPVFLVPPPEITASEVESTELSTATFTSTQSPLOKLFARKMKILGTQILRGI 59
Db 11 LGSASAHGTPGSTGPELNTSVH-----PINGSPOYQK---AKLOVLGAIGILNAA 60

QY 60 MTFSGVIEFLTKPKY----PRPFIFLGGYPPMGWSVLFINSGAPFLIAVKRKTETLII 115
Db 61 MILALGV-FLGSLQPYHFQKHFFFTFTGYTGVAGVAFVCSSTGTSVVGAKPRTWIIQ 119

QY 116 LSRIMFLSALGAAGIILLTFGLIQONI--CGYSHQNSQ----CKATVTLFGLIIT 169
Db 120 NSFGMNIASATALVGTAFLSLNIAVNIQSLRSC---HSSSESPDLNCNMGSIQMSVL 176

QY 170 LMTFSIIEFLSILPFSILGCHSEDCDCEQ 198
Db 177 LLITLLELCVTISTIAIWMNCNCSNR 205

```

[illegible]

RESULT 4
US-08-707-340-4
; Sequence 4, Application US/08707340
; Patent No. 5705615
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,340
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,492
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/675,648
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BIH94-03A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-707-340-4

Query Match 17.7%; Score 181.5; DB 1; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 38; Mismatches 87; Indels 25; Gaps 8;
QY 1 MDSTAH-SPVFLVFPPEITASEYESTELSATFTSTQSPLOKLPARKMKILGTIQLFGI 59
Db 11 LGSASAHGTPGSETGPEELNTSVYH-----PINGSPTYQK---AKQLVLGAIIQILNAA 60
QY 60 MTFSEGVIFLTLKPY----PRFPFIILSGYPFGWSVLFINSAGFLIAVKRKTETLII 115
Db 61 MILALGV-FLGSLQYPHFKHFFFTYGYPIWGAVFFCSSLVSVVAGIKPTRTWIQ 119
QY 116 LSRIMNLSALGAIIGILLTGFILDQNYI--CGYSHQNSQ----CKAVTVLFLGLIT 169
Db 120 NSFQMIASATIALVGTAFSLNIAVNIQSLRSC---HSSSESPDLNMGNSISNGMVS 176
QY 170 LMTFSIIEFLPSILCHSEDCDEQ 198
Db 177 LLTLLELCVTISTIAMWCNANCCNSRE 205

RESULT 5

US-08-994-578-4

; Sequence 4, Application US/08994578

US-08-916-902A-3
; Sequence 3, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 561639
US-08-916-902A-3

Query Match 17.7%; Score 181.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 38; Mismatches 87; Indels 25; Gaps 8;
QY 1 MDSTAH-SPVFLVFPPEITASEYESTELSATFTSTQSPLOKLPARKMKILGTIQLFGI 59
Db 11 LGSASAHGTPGSETGPEELNTSVYH-----PINGSPTYQK---AKQLVLGAIIQILNAA 60
QY 60 MTFSEGVIFLTLKPY----PRFPFIILSGYPFGWSVLFINSAGFLIAVKRKTETLII 115
Db 61 MILALGV-FLGSLQYPHFKHFFFTYGYPIWGAVFFCSSLVSVVAGIKPTRTWIQ 119
QY 116 LSRIMNLSALGAIIGILLTGFILDQNYI--CGYSHQNSQ----CKAVTVLFLGLIT 169
Db 120 NSFQMIASATIALVGTAFSLNIAVNIQSLRSC---HSSSESPDLNMGNSISNGMVS 176
QY 170 LMTFSIIEFLPSILCHSEDCDEQ 198
Db 177 LLTLLELCVTISTIAMWCNANCCNSRE 205

RESULT 6

US-08-994-578-4

; Sequence 4, Application US/08994578


```

; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: beta subunit
; US-07-869-933-29

Query Match 15.4%; Score 158; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 3e-10;
Matches 45; Conservative 28; Mismatches 72; Indels 16; Gaps 3;

QY 36 QSPLOKLFARKMKILGTQILFGIMTFSGVIFLTLKP--YPRFPFIPLSGYPFWGVS 93
Db 47 QQTWQSFLLKKELEFGVTQVLGLICFCFVVCSTLQTSDFDDEVLVLLYRAGYPPFWGAV 106

QY 94 LFINSGAFLLIYAVKRKTTETLILSRIMNLSALGAIAGIILLTFGLDQNYICGYSHQN 153
Db 107 LFLVLSGFLSIMSERKNTLYVRGSLGANIVSSIAAGLGIAILLNLNSNAYM----- 159

QY 154 SOCKAVT-----VLFGILITLMTFSLIELFSLPSIL 187
Db 160 NYCKDITDDGCFVTSFTELVLMLLFTILAFCSAVLLII 200

RESULT 9
US-07-869-933-33
; Sequence 33, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: beta subunit
; US-07-869-933-29

Query Match 15.4%; Score 158; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 3e-10;
Matches 45; Conservative 28; Mismatches 72; Indels 16; Gaps 3;

QY 36 QSPLOKLFARKMKILGTQILFGIMTFSGVIFLTLKP--YPRFPFIPLSGYPFWGVS 93
Db 47 QQTWQSFLLKKELEFGVTQVLGLICFCFVVCSTLQTSDFDDEVLVLLYRAGYPPFWGAV 106

QY 94 LFINSGAFLLIYAVKRKTTETLILSRIMNLSALGAIAGIILLTFGLDQNYICGYSHQN 153
Db 107 LFLVLSGFLSIMSERKNTLYVRGSLGANIVSSIAAGLGIAILLNLNSNAYM----- 159

QY 154 SOCKAVT-----VLFGILITLMTFSLIELFSLPSIL 187
Db 160 NYCKDITDDGCFVTSFTELVLMLLFTILAFCSAVLLII 200

RESULT 9
US-07-869-933-33
; Sequence 33, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
```

```

; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: rat
; STRAIN: FCRI beta subunit
; US-07-869-933-33

Query Match 15.4%; Score 158; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 3e-10;
Matches 45; Conservative 28; Mismatches 72; Indels 16; Gaps 3;

QY 36 QSPLOKLFARKMKILGTQILFGIMTFSGVIFLTLKP--YPRFPFIPLSGYPFWGVS 93
Db 47 QQTWQSFLLKKELEFGVTQVLGLICFCFVVCSTLQTSDFDDEVLVLLYRAGYPPFWGAV 106

QY 94 LFINSGAFLLIYAVKRKTTETLILSRIMNLSALGAIAGIILLTFGLDQNYICGYSHQN 153
Db 107 LFLVLSGFLSIMSERKNTLYVRGSLGANIVSSIAAGLGIAILLNLNSNAYM----- 159

QY 154 SOCKAVT-----VLFGILITLMTFSLIELFSLPSIL 187
Db 160 NYCKDITDDGCFVTSFTELVLMLLFTILAFCSAVLLII 200

RESULT 10
US-08-201-879A-4
; Sequence 4, Application US/08201879A
; Patent No. 5807988
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; APPLICANT: JOUVIN, Marie-Helene
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,879A
; FILING DATE: 24-FEB-1994
```


GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
IMMUNOGLOBULIN
TITLE OF INVENTION: IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22113-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 246 amino acids

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 11:02:21 ; Search time 11.95 Seconds
(without alignments)
648.026 Million cell updates/sec

Title: US-09-821-821-2

Perfect score: 1027

Sequence: 1 MDSSTAHPVFLVFPPEITA.....SLPFSILGSHSDCDCEQCC 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Length | DB ID | Description |
|------------|-------|---------|--------|-------|-------------|
| 1 | 173 | 16.8 | 291 | 1 | CD20_MOUSE |
| 2 | 169 | 16.5 | 297 | 1 | CD20_HUMAN |
| 3 | 158 | 15.4 | 243 | 1 | FCEB_RAT |
| 4 | 157.5 | 15.3 | 235 | 1 | FCEB_MOUSE |
| 5 | 123.5 | 12.0 | 244 | 1 | FCEB_HUMAN |
| 6 | 98.5 | 9.6 | 458 | 1 | TCR_STRAG |
| 7 | 94 | 9.2 | 408 | 1 | Y917_AQUAE |
| 8 | 91.5 | 8.9 | 458 | 1 | TCR_EACST |
| 9 | 91.5 | 8.9 | 458 | 1 | TCR_STREN |
| 10 | 90 | 8.8 | 446 | 1 | NU4M_DROME |
| 11 | 87.5 | 8.5 | 656 | 1 | YAN9_SCHPO |
| 12 | 87.5 | 8.5 | 956 | 1 | MTN2_MOUSE |
| 13 | 86.5 | 8.4 | 333 | 1 | NU2M_APTLI |
| 14 | 86.5 | 8.4 | 666 | 1 | NU5M_CHOCR |
| 15 | 86 | 8.4 | 300 | 1 | Y281_BUCAP |
| 16 | 85 | 8.3 | 382 | 1 | RFC_SHIFL |
| 17 | 85 | 8.3 | 576 | 1 | DSBD_PASMU |
| 18 | 85 | 8.3 | 652 | 1 | NU5M_PODAN |
| 19 | 85 | 8.3 | 1030 | 1 | VPP1_CABEL |
| 20 | 84.5 | 8.2 | 447 | 1 | NU4M_ANOGA |
| 21 | 84 | 8.2 | 198 | 1 | COP_CLOPE |
| 22 | 84 | 8.2 | 216 | 1 | SSPN_MOUSE |
| 23 | 84 | 8.2 | 463 | 1 | YS97_CAEEL |
| 24 | 83.5 | 8.1 | 458 | 1 | NAH1_PIG |
| 25 | 83.5 | 8.1 | 818 | 1 | NAH1_PIG |
| 26 | 83 | 8.1 | 880 | 1 | YE21_ARCFU |
| 27 | 82.5 | 8.0 | 230 | 1 | Y920_HELPY |
| 28 | 82.5 | 8.0 | 409 | 1 | NU4M_ASCSU |
| 29 | 82.5 | 8.0 | 525 | 1 | COX1_CAEEL |
| 30 | 82.5 | 8.0 | 590 | 1 | CAN1_YEAST |
| 31 | 82.5 | 8.0 | 817 | 1 | NAH1_BOVIN |
| 32 | 82 | 8.0 | 657 | 1 | NU5M_EMENI |
| 33 | 81.5 | 7.9 | 230 | 1 | Y920_HELPY |

RESULT 1
CD20_MOUSE
ID CD20_MOUSE STANDARD; PRT; 291 AA.
AC P19437;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE B-cell surface protein CD20 homolog (B-cell differentiation antigen LY-44).
DE MS4A1 OR CD20 OR LY-44 OR MS4A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=89067519; PubMed=2461992;
RA Tedder T.F., Klejman G., Distech C.M., Adler D.A., Schlossman S.F., Saito H.;
RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation antigen, homologous to the human B1 (CD20) antigen, and localization of the gene to chromosome 19.";
RL J. Immunol. 141:4388-4394(1988).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Ouackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult D., Hofmann M., Hume D.A., Kamliya M., Lee N.H., Gustincich S., Hill D., Hofmann M., Mazarrelli J., Mombaerts P., Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakamoto N., Nordone P., Ring B., Ringwald M., Ruggieri J., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F., Suzuki H., Tovo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
RL Nature 409:685-690(2001).
RT "Functional annotation of a full-length mouse cDNA collection.";
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

34 81.5 7.9 252 1 MPUL_DROME Q9mw8 drosophila
35 81.5 7.9 525 1 COX1_ASCSU P24861 ascaris suu
36 81.5 7.9 816 1 NAH1_RABIT P23791 oryctolagus
37 81 7.9 291 1 YC38_PORPU P51321 porphyra pu
38 81 7.9 298 1 Y680_HAEIN Q57389 haemophilus
39 81 7.9 324 1 NUIM_DROYA P07710 drosophila
40 81 7.9 397 1 YLH2_CAEEL P34356 caenorhabdi
41 80.5 7.8 365 1 CYB_ASCSU P24878 ascaris suu
42 80.5 7.8 373 1 CKR2_RAT O55193 rattus norv
43 80.5 7.8 448 1 NU4M_ANOOU P33511 anopheles q
44 80.5 7.8 815 1 NAH1_HUMAN P19634 homo sapien
45 80 7.8 371 1 NUIM_NEUCR P08774 neurospora

ALIGNMENTS

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M62541; AAA37394.1; -
EMBL; AK017903; BAB30996.1; -
PIR; A30558; A30558.
MGD; MGI:88321; Ms4a2.
B-cell; Transmembrane; Phosphorylation.
FT DOMAIN 1 44 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT DOMAIN 204 291 CYTOPLASMIC (POTENTIAL).
SEQUENCE 291 AA; 31958 MW; D478ECD2C5C16FC CRC64;

Query Match 16.8%; Score 173; DB 1; Length 291;
Best Local Similarity 29.4%; Pred. No. 2.7e-07;
Matches 53; Conservative 29; Mismatches 64; Indels 34; Gaps 7;

QY 32 NFSTOSPLQKLFARKMKILGTITQILFGIMTFSGVIFLFTLLKYPFRPPFFILS-GYFEW 90
DB 27 TSSLVGPQTSFMRKSKALGAVQIMNGLFHITLGLG-----LMIPTGVFAPICLSVWYPLW 82
QY 91 GSVLFINSGLFIIVKRRKTTETLILSR-IMNFLSALGAIAGIIL-----LTFGILD 142
DB 83 GGIMYIISGSLAAAEKTSKSLVKAVINSSLSLFAAISGILSMDILNMTLSHFLK 142
QY 143 QN-----YIC-----GYSHONS-----OCKAVTVLFLGILITLMTFSIIELFIS 181
DB 143 MRRLIQTSPKPYVDIYDCEPNSSEKNSPSTQYCSIQSVFLGILSAMLISAFQKLVT 202

RESULT 2
ID CD20_HUMAN STANDARD; PRT; 297 AA.
AC P11836; P08984; Q13963;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen CD20 (B-lymphocyte surface antigen B1) (Leu-16) (Bp35).
GN MS4A1 OR CD20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88124792; PubMed=2448768;
RA Stamenkovic I., Seed B.;
RT "Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1, Bp35), a type III integral membrane protein.";
RL J. Exp. Med. 167:1975-1980(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88124792; PubMed=2448768;
RA Tedder T.F., Streuli M., Schlossman S.F., Saito H.;
RT "Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes";
RL Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89176281; PubMed=2466899;
RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
RT "Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (B1).";
RL J. Immunol. 142:2560-2568(1989).
RN [4]
RP SEQUENCE FROM N.A.

RX MEDLINE=88283639; PubMed=2456210;
RA Einfeld D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;
RT "Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains.";
RL EMBO J. 7:711-717(1988).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN KINASE(S).
CC -!- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD20 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd20.htm".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X12530; CAA31046.1; -
EMBL; M27394; AAA35581.1; -
EMBL; M27395; -; NOT_ANNOTATED_CDS.
EMBL; L23419; AAA88911.1; -
EMBL; L23415; AAA88911.1; JOINED.
EMBL; L23416; AAA88911.1; JOINED.
EMBL; L23417; AAA88911.1; JOINED.
EMBL; L27203; CAA30179.1; -
EMBL; X07204; CAA30180.1; -
PIR; A27400; A27400.
PIR; J10042; J10042.
PIR; A30586; A30586.
PIR; S00387; S00387.
MIM; 112210; -
KW B-cell; Transmembrane; Phosphorylation.
FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT DOMAIN 210 297 CYTOPLASMIC (POTENTIAL).
FT DISULFID 81 167 PROBABLE.
FT DISULFID 167 183 PROBABLE.
FT DISULFID 111 220 PROBABLE.
FT CONFLICT 13 13 P -> L (IN REF. 4).
FT CONFLICT 71 71 M -> I (IN REF. 3).
SEQUENCE 297 AA; 33077 MW; AC5420F8B626BDD1 CRC64;

Query Match 16.5%; Score 169; DB 1; Length 297;
Best Local Similarity 30.3%; Pred. No. 5.9e-07;
Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;
QY 34 STOSPLQKLFARKMKILGTITQILFGIMTFSGVIFLFTLLKYPFR-PFIFLSGYPFWGS 92
DB 36 SLVGPQTSFMRKSKALGAVQIMNGLFHITLGLG-----LMIPAGIYAPICVWYPLWGG 91
QY 93 VLFINSGLFIIVKRRKTTETLILSRIMNFLSALGAIAGIIL-----T 136
DB 92 IMYIISGSLAAAEKTSKSLVKAVINSSLSLFAAISGILSMDILNMTLSHFLKMS 151
QY 137 FGFILDQN-YICG-----SHONS-----OCKAVTVLFLGILITLMTFSII-ELFIS 181
DB 152 LNFIRAHPTPIYINYNCEPANPSEKNSPSTQYCSIQSVFLGILSAMLISAFQELVIA 209
RESULT 3
ID FCEB_RAT STANDARD; PRT; 243 AA.
AC P13386;
DT 01-JAN-1990 (Rel. 13, Created)

Db 117 YLVRSUGANIVSIAAGTGIAMLIL-----NLTNFAYMNN-CKNVTEDDGCVFASFT 169
QY 165 GILITLMTFSIIELIFS 181
Db 170 TELVLMFLTLILAFCS 186

RESULT 5
FCBE_HUMAN STANDARD; PRT; 244 AA.
AC Q01362;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCERI)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316966; PubMed=1535625;
RA Kuester H., Zhang L., Brini A.T., Macglashan D.W., Kinet J.-P.;
RT "The gene and cDNA for the human high affinity immunoglobulin E
RT receptor beta chain and expression of the complete human receptor.";
RL J. Biol. Chem. 267:12782-12787(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339505; PubMed=1386024;
RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
RT "Determination of the sequence coding for the beta subunit of the
RT human high-affinity Ige receptor.";
RL FEBS Lett. 302:161-165(1992).
RN [3]
RP VARIANT GLY-237.
RX MEDLINE=96414302; PubMed=8817330;
RA Hill M.R., Cookson W.O.;
RT "A new variant of the beta subunit of the high-affinity receptor for
RT immunoglobulin E (Fc epsilon RI-beta E237G): associations with
RT measures of atopy and bronchial hyper-responsiveness.";
RL Hum. Mol. Genet. 5:959-962(1996).
RN [4]
RP VARIANT GLY-237.
RX MEDLINE=96440420; PubMed=8842731;
RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RA Hopkin J.;
RT "Association between atopic asthma and a coding variant of Fc-epsilon-
RT RI-beta in a Japanese population.";
RL Hum. Mol. Genet. 5:1129-1130(1996).
RN [5]
RP ERRATUM.
RX MEDLINE=97123518; PubMed=8968765;
RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RA Hopkin J.;
RL Hum. Mol. Genet. 5:2068-2068(1996).
CC [-] FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC [-] SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC [-] SUBCELLULAR LOCATION: Integral membrane protein.
CC [-] TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
CC BASOPHILS.
CC [-] POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE PRESENT IN ABOUT
CC 5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
CC SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
CC [-] SIMILARITY: BELONGS TO THE MS4A FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D10583; BAA01440.1; -;
DR EMBL; M89796; AAA60269.1; -;
DR PIR; S21154; S21154.
DR PIR; A42806; A42806.
DR MIN; 147138; -;
KW Ige-binding protein; Receptor; Transmembrane; Polymorphism.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 79 POTENTIAL.
FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 117 POTENTIAL.
FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 150 POTENTIAL.
FT DOMAIN 151 180 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 181 200 POTENTIAL.
FT DOMAIN 201 244 CYTOPLASMIC (POTENTIAL).
FT VARIANT 237 237 E-> G.
SQ SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;
/FTID=VAR_003965.
Query Match 12.08; Score 123.5; DB 1; Length 244;
Best Local Similarity 25.6%; Pred. No. 0.0023;
Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;
QY 9 PVFLVFPPEITASEVSETELSATFTSTQSIQ---KLFARKMKILGTIQLFGIMTFSFG 65
Db 21 PAFEVL--EISPFQVSSGRLKLS--ASSPPLHTWLVTKQEFLGVQILTANICICFG 76
QY 66 VIFLFTLLKPYRPPFI--FLSGYPFWGVSFLFINSGLFIAVKRKTTETLILSRINFL 123
Db 77 TVCVSLVDISHIEGDFSSFRAGYPFWGAIFFSIGMLSIISERRNATYLVRSIGANTA 136
QY 124 SALGAIGAIIITGFIIDQNYICGYSHONS----QCKAVTVLFLGLITLMTF-SIIE 177
Db 137 SSIAGGTGITIL----IINLKKSLAYIIHSCQKFFETKCFMASFSTEIVVMFLTLILG 192
QY 178 LFISLPFSILGCHSE 192
Db 193 LGSVSLTICGAGEE 207
RESULT 6
TCR_STRAG STANDARD; PRT; 458 AA.
AC P13924;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Tetracycline resistance protein.
GN TET.
OS Streptococcus agalactiae.
OC Plasmid pW158.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90016790; PubMed=2677995;
RA van der Lelie D., Bron S., Venema G., Oskam L.;
RT "Similarity of minus origins of replication and flanking open reading
RT frames of plasmids pUB110, pTB913 and pW158.";
RL Nucleic Acids Res. 17:7283-7294(1989).
CC [-] FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE


```

DR InterPro: IPR000260; Oxidored_q5_N.
DR Pfam: PF00361; oxidored_q1; 1.
DR Pfam: PF01059; oxidored_q5_N; 1.
DR PRINTS: PR01437; NUOXDRDTASE4.
KW Oxidoreductase; NAD; Ubiquitine; Mitochondrion.
FT CONFLICT 161 161 S-> V (IN REF. 1; AAC47819).
SQ SEQUENCE 446 AA; 51371 MW; 8AB4A98563238BB2 CRC64;

Query Match      8.8%; Score 90; DB 1; Length 446;
Best Local Similarity 23.6%; Pred. No. 2;
Matches 34; Conservative 25; Mismatches 45; Indels 40; Gaps 6;

QY 66 VIFLFTLLKPYRFRFFILSGYPFW--GSLVLFINSGLFIATVRKKTTETLIILSRIMNKL 123
   : || | : || | : : || | : || | : || | : || | : || | : || | : || | :
DB 5 IFELLFLI-----PFCFINNN-YMWQIMMFFISFIFL-----LMNNFNMTW 45
   : || | : || | : || | : || | : || | : || | : || | : || | : || | :
QY 124 SALCAGTAGIILLTFGFILDONVICGY-----SHQNSOCKAVTVLFGILITLTMTFS 174
   : || | : || | : || | : || | : || | : || | : || | : || | : || | :
DB 46 SEISYELGCDMLSYGLILLSLWICSLMLASEMKNNKYKNFLNLIIILLLLLLTLTFS 105
   : || | : || | : || | : || | : || | : || | : || | : || | : || | :
QY 175 IIELF-----ISLPFSILG 188
   : || | : || | : || | : || | : || | : || | : || | : || | : || | :
DB 106 SMSLEFMFYLFESSLIPTLFILIG 129
   : || | : || | : || | : || | : || | : || | : || | : || | : || | :

RESULT 11
ID YAN9_SCHPO STANDARD; PRT; 656 AA.
AC Q10074;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 73.1 kDa protein C3HL.09c in chromosome I.
GN SPAC3H1.09C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_taxid=4896;
RX NCBI
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentiles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO YEAST YKL146W AND YNL101W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z68144; CAA92262.1; -
DR InterPro: IPR002422; AA_rel_permease_2.
DR Pfam: PF01490; Aa_trans; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 389 409 POTENTIAL.
FT TRANSMEM 419 439 POTENTIAL.
FT TRANSMEM 457 477 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 537 557 POTENTIAL.
FT TRANSMEM 578 598 POTENTIAL.
FT TRANSMEM 601 621 POTENTIAL.
FT TRANSMEM 636 656 POTENTIAL.
SQ SEQUENCE 656 AA; 73060 MW; 7FDIE8F649EDB3D2 CRC64;
```

Query Match 8.5%; Score 87.5; DB 1; Length 656;
 Best Local Similarity 22.7%; Pred. NO. 4.6;
 Matches 54; Conservative 34; Mismatches 75; Indels 75; Gaps 10;

QY 11 FLVPPPEITASEVESFELSATFTSQSPLOKLFARKMKILGTIQLF----- 57
 DB 400 FLVFPV--LSLVKRSKLSATA-----LIADVILLGLLYFYFDVITLTKGIAD 448
 QY 58 -----GIMTSGFVIFLF-----TLLKP--YPRPFIFLSGYPFGWSVLFIN 97
 DB 449 VAMENKTDPSLFGVAITFYEGICILIPQEQMARPKNLK-----LLGVMAAISLLFIS 504
 QY 98 SCAFLIAVKRKTEFLIILSR-----IMNPLSALGAIAGIILTFG--FILDQNYICG 148
 DB 505 IGLLSYAAGSGVKVKTWVILNMPSTFTWIIQFLYAIALLSTPLQLPFAIAIEQIGFTR 564
 QY 149 YSHQNSQCK-----AVTVLFGLITLMTSLIELFSLPFSILGCHSEDCDCQCC 200
 DB 565 SGRNRKIKWRKNYLRVLIVLAILISWAGSRDLFVSMGVS-----CC 610

RESULT 12
 MTN2_MOUSE STANDARD; PRT; 956 AA.

AC 008746;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 GN Matrilin-2 precursor.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Limb;
 RX MEDLINE=97238863; Pubmed=9083061;
 RA Deak F., Piecha D., Bachrati C., Paulsson M., Kiss I.;
 RT Primary structure and expression of matrilin-2, the closest relative
 RT of cartilage matrix protein within the von Willebrand factor type A-
 RT like module superfamily.
 RL J. Biol. Chem. 272:9268-9274 (1997).
 CC -1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
 CC CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
 CC OSTEOBLAST CELL LINES.
 CC -1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 VWFA DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U69262; AAC53163.1; .
 DR HSSP; P35555; IEMN.
 DR MGD; MGI:109613; Matn2.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF-Ca.
 DR InterPro; IPR02035; VWFA.
 DR Pfam; PF00008; EGF; 10.
 DR Pfam; PF00092; vwa; 2.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00181; EGF; 10.
 DR SMART; SM00327; VWA; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 7.

DR PROSITE; PS01186; EGF_2; 9.
 DR PROSITE; PS50234; VWFA_2.
 KW EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil.
 FT SIGNAL 1 POTENTIAL.
 FT CHAIN 24 956 MATRILIN-2.
 FT DOMAIN 57 232 VWFA 1.
 FT DOMAIN 238 278 EGF-LIKE 1.
 FT DOMAIN 279 319 EGF-LIKE 2.
 FT DOMAIN 320 360 EGF-LIKE 3.
 FT DOMAIN 361 401 EGF-LIKE 4.
 FT DOMAIN 402 442 EGF-LIKE 5.
 FT DOMAIN 443 483 EGF-LIKE 6.
 FT DOMAIN 484 524 EGF-LIKE 7.
 FT DOMAIN 525 565 EGF-LIKE 8.
 FT DOMAIN 566 606 EGF-LIKE 9.
 FT DOMAIN 607 647 EGF-LIKE 10.
 FT DOMAIN 655 830 VWFA 2.
 FT DOMAIN 917 955 COILED COIL (POTENTIAL).
 FT DISULFID 242 253 BY SIMILARITY.
 FT DISULFID 249 262 BY SIMILARITY.
 FT DISULFID 264 277 BY SIMILARITY.
 FT DISULFID 283 294 BY SIMILARITY.
 FT DISULFID 290 303 BY SIMILARITY.
 FT DISULFID 305 318 BY SIMILARITY.
 FT DISULFID 324 335 BY SIMILARITY.
 FT DISULFID 331 344 BY SIMILARITY.
 FT DISULFID 346 359 BY SIMILARITY.
 FT DISULFID 365 376 BY SIMILARITY.
 FT DISULFID 372 385 BY SIMILARITY.
 FT DISULFID 387 400 BY SIMILARITY.
 FT DISULFID 406 417 BY SIMILARITY.
 FT DISULFID 413 426 BY SIMILARITY.
 FT DISULFID 428 441 BY SIMILARITY.
 FT DISULFID 447 458 BY SIMILARITY.
 FT DISULFID 454 467 BY SIMILARITY.
 FT DISULFID 469 482 BY SIMILARITY.
 FT DISULFID 488 499 BY SIMILARITY.
 FT DISULFID 495 508 BY SIMILARITY.
 FT DISULFID 510 523 BY SIMILARITY.
 FT DISULFID 529 540 BY SIMILARITY.
 FT DISULFID 536 549 BY SIMILARITY.
 FT DISULFID 551 564 BY SIMILARITY.
 FT DISULFID 570 581 BY SIMILARITY.
 FT DISULFID 577 590 BY SIMILARITY.
 FT DISULFID 592 605 BY SIMILARITY.
 FT DISULFID 611 622 BY SIMILARITY.
 FT DISULFID 618 631 BY SIMILARITY.
 FT DISULFID 633 646 BY SIMILARITY.
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 956 AA; 106779 MW; 3B4C22770B6E6EE CRC64;

Query Match 8.5%; Score 87.5; DB 1; Length 956;
 Best Local Similarity 23.2%; Pred. NO. 6.4;
 Matches 55; Conservative 31; Mismatches 76; Indels 75; Gaps 13;

QY 11 FLVFPPEIT--ASEYEST--ELSATFTSQSPLOKLFARKMKILGTIQLFQIMT--- 61
 DB 86 FLDIGPDVTRVGLLYGSGTVKNEFSLKTFKKSEVERA-VKMRHLST-----GMTGLA 139
 QY 62 --FSFGVIF-----LFTLLKPYPRPFIFLSGYPFGWSVLFINSGAFLIAVKRKTTETLI 114
 DB 140 IOYALNIAFSEAEGARPLRENVPRIIMIVTDGRP-----QDSVAEVAKARNTGLI 191
 QY 115 ILSRI-----MNFLSALGA--TAGIILLTFGLIDONYICGSHQNSQCKAVTVLFLGLI 168
 DB 192 FAIGVGQVDLNTLKAIGSEPHKDHVFLVANF-----SQIESLTSVFQNKLC 237
 QY 169 TLMTFSIIE-----LFTSLPFS-----ILG-----CHSEDCDCQCC 200
 DB 238 TVHMCVLEHNCACHECLNTPGSYICKCKQGYILSTDQKTCRIQDLCCATEDHGCEQLC 294

```

RESULT 13
NU2M,APILI
ID NU2M,APILI STANDARD: PRT: 333 AA.
AC P34349; P92496; P92497; P92502; P92886; P92907; Q33785; Q33786;
AC Q33787; Q33788; Q36885;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3).
GN ND2.
OS Apis mellifera ligustica (Common honeybee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7469;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thorax;
RX MEDLINE=93114603; PubMed=84117993;
RA Crozier R.H., Crozier Y.C.;
RT "The mitochondrial genome of the honeybee Apis mellifera: complete
RT sequence and genome organization.";
RL Genetics 133:97-117(1993).
RN [2]
RP SEQUENCE OF 1-211 FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RX MEDLINE=96364025; PubMed=8744768;
RA Arias M.C., Sheppard W.S.;
RT "Molecular phylogenetics of honey bee subspecies (Apis mellifera L.)
RT inferred from mitochondrial DNA sequence.";
RL Mol. Phylogenet. Evol. 5:557-566(1996).
RN [3]
RP SEQUENCE OF 60-307 FROM N.A.
RC STRAIN=HAPLOTYPES 1 TO 11;
RA Koulianos S., Crozier R.H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -! SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -! SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L06178; AAB96798.1; -
DR EMBL; U35743; AAB38223.1; -
DR EMBL; U35744; AAB38224.1; -
DR EMBL; U35745; AAB38225.1; -
DR EMBL; U35746; AAB38226.1; -
DR EMBL; U35747; AAB38227.1; -
DR EMBL; U35748; AAD12744.1; -
DR EMBL; U35749; AAD12745.1; -
DR EMBL; U35750; AAB38228.1; -
DR EMBL; U35751; AAB38229.1; -
DR EMBL; U35752; AAB38230.1; -
DR EMBL; U35753; AAB38231.1; -
DR EMBL; U35754; AAB38232.1; -
DR EMBL; U35755; AAB38233.1; -
DR EMBL; U35756; AAB38234.1; -
DR EMBL; U35757; AAB38235.1; -
DR EMBL; U35758; AAB38236.1; -
DR EMBL; U35759; AAB38237.1; -
DR EMBL; U35760; AAB38238.1; -
DR EMBL; U35761; AAB38239.1; -
DR EMBL; U35762; AAB38241.1; -
DR EMBL; U35763; AAB38242.1; -

```

```

DR EMBL; U35764; AAB38243.1; -
DR EMBL; U35765; AAB38244.1; -
DR EMBL; U72279; AAB41177.1; -
DR EMBL; U72280; AAB41178.1; -
DR EMBL; U72281; AAB41179.1; -
DR EMBL; U72282; AAB41180.1; -
DR EMBL; U72283; AAB41181.1; -
DR EMBL; U72284; AAB41182.1; -
DR EMBL; U72285; AAB41183.1; -
DR EMBL; U72286; AAB41184.1; -
DR EMBL; U72287; AAB41185.1; -
DR EMBL; U72288; AAB41186.1; -
DR EMBL; U72289; AAB41187.1; -
DR InterPro: IPR001750; Oxidored_q1.
DR Pfam: PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT VARIANT 14 14 F -> L (IN SSP. ADANSONII, SSP. CAPENSIS,
FT VARIANT 62 62 F -> L (IN SSP. ADANSONII, SSP. CAPENSIS,
FT VARIANT 84 84 INTERMISSA, SSP. IBERICA, SSP.
FT VARIANT 138 138 MACEDONICA, SSP. LAMARCKII, SSP. SSP.
FT VARIANT 142 142 SSP. SAHARIENSIS, SSP. SCUTELLATA AND
FT VARIANT 158 158 SSP. SICULA).
FT VARIANT 203 203 V -> M (IN HAPLOTYPE 9).
FT VARIANT 262 262 T -> I (IN HAPLOTYPES 3, 5, 6, 7 AND 10).
FT VARIANT 333 333 V -> I (IN HAPLOTYPE 4, MELLII AND
SQ SEQUENCE 333 AA; 40879 MW; 866969F11529907F CRC64;
MELLII2).
Query Match 8.4%; Score 86.5; DB 1; Length 333;
Best Local Similarity 25.1%; Pred. No. 2.9; Indels 25; Gaps 9;
Matches 48; Conservative 33; Mismatches 85;
QY 8 SPVFLVPPPEITASEYESTELSATFTTSQSPQLKLFARKMKILGTIQLFGIMTFSGV- 66
DB 66 SSIFLFFMIIVLSSISFTKTDTFNFM----VQMMFFLK---IGTFPPHFW-MIYSYEMM 117
QY 67 ----IFLF-TLLKPYPRFFFLSGYPFWGSLVFNISGAFI--TAVRKKTETLIIISRI 119
DB 118 NWKQIFLMSTLFIPIFYMMVSMKINSWTLFLTNSLIYSFYANKFYTLKKLACSTI 177
QY 120 MN-----FLSALGA---IAGIILLTFGLDQNYICGYSHONSQCKAVTVLFLGILITLM 171
DB 178 FNSFYFIFILELKNMFAMIIYSFNYELLISFLNKNQNFNFNFYFNKYQMYTFLTLM 237
QY 172 -TFSIIELEFIS 181
DB 238 FNYSMYPIFLS 248
RESULT 14
NU5M,CHOCH
ID NU5M,CHOCH STANDARD: PRT: 666 AA.
AC P48920; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NAD5.
OS Chondrus crispus (Carragheen).
OG Mitochondrion.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartiniaceae;
OC Chondrus.
OX NCBI_TaxID=2769;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Apices;
RX MEDLINE=95341681; PubMed=7616569;
RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienenberger J.M.,

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 11:02:16 ; Search time 19.84 Seconds
(without alignments)
968.642 Million cell updates/sec

Title: US-09-821-821-2

Perfect score: 1027

Sequence: 1 MDSSTAHSPVFLVFPPEITA.....SLPFSILGCHSEDCDCQCC 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 1027 | 100.0 | 200 | 2 | JC7585 |
| 2 | 181.5 | 17.7 | 214 | 2 | I59258 |
| 3 | 173 | 16.8 | 291 | 2 | A30558 |
| 4 | 169 | 16.5 | 297 | 1 | A30586 |
| 5 | 158 | 15.4 | 243 | 2 | A31231 |
| 6 | 157.5 | 15.3 | 235 | 2 | B34342 |
| 7 | 123.5 | 12.0 | 244 | 2 | A42806 |
| 8 | 106 | 10.3 | 307 | 2 | C86844 |
| 9 | 103 | 10.0 | 455 | 2 | T16070 |
| 10 | 100 | 9.7 | 152 | 2 | JH0751 |
| 11 | 96 | 9.3 | 387 | 2 | E81306 |
| 12 | 95.5 | 9.3 | 537 | 2 | T07892 |
| 13 | 95 | 9.3 | 165 | 2 | F69208 |
| 14 | 94.5 | 9.2 | 991 | 2 | B71315 |
| 15 | 94 | 9.2 | 408 | 2 | C70379 |
| 16 | 94 | 9.2 | 439 | 2 | E95155 |
| 17 | 94 | 9.2 | 439 | 2 | H98021 |
| 18 | 93.5 | 9.1 | 542 | 2 | T01124 |
| 19 | 92.5 | 9.0 | 533 | 2 | T07894 |
| 20 | 91.5 | 8.9 | 458 | 1 | YBBSRT |
| 21 | 91.5 | 8.9 | 458 | 1 | YBBSU6 |
| 22 | 91.5 | 8.9 | 458 | 1 | YTSOG |
| 23 | 91.5 | 8.9 | 458 | 1 | YTSOG |
| 24 | 91 | 8.9 | 288 | 2 | A05235 |
| 25 | 91 | 8.9 | 1431 | 2 | T22748 |
| 26 | 90.5 | 8.8 | 457 | 2 | T37205 |
| 27 | 90 | 8.8 | 446 | 2 | S01187 |
| 28 | 89.5 | 8.7 | 1400 | 2 | A71514 |
| 29 | 89 | 8.7 | 244 | 2 | H70193 |

| | | | | | |
|----|------|-----|------|---|--------|
| 30 | 89 | 8.7 | 353 | 2 | T42971 |
| 31 | 88.5 | 8.6 | 275 | 2 | H96981 |
| 32 | 88.5 | 8.6 | 538 | 2 | T05714 |
| 33 | 88.5 | 8.6 | 538 | 2 | T07164 |
| 34 | 88.5 | 8.6 | 540 | 2 | T07604 |
| 35 | 88.5 | 8.6 | 694 | 2 | E69143 |
| 36 | 88 | 8.6 | 485 | 2 | E98261 |
| 37 | 88 | 8.6 | 485 | 2 | AD3023 |
| 38 | 87.5 | 8.5 | 319 | 2 | H69160 |
| 39 | 87.5 | 8.5 | 369 | 2 | H90587 |
| 40 | 87.5 | 8.5 | 448 | 2 | H97008 |
| 41 | 87.5 | 8.5 | 656 | 2 | T38741 |
| 42 | 87.5 | 8.5 | 1227 | 2 | T20370 |
| 43 | 87.5 | 8.5 | 1400 | 2 | A81672 |
| 44 | 87 | 8.5 | 141 | 2 | H89831 |
| 45 | 87 | 8.5 | 226 | 2 | F84086 |

ALIGNMENTS

RESULT 1
JC7585
testis expressed transmembrane-4 protein, TETM4 - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7585
R:Hulett, M.D.; Pagler, E.; Hornby, J.R.; Hogarth, P.M.; Eyre, H.J.; Baker, E.; Crawford
Biochem. Biophys. Res. Commun. 280, 374-379, 2001
A:Title: Isolation, tissue distribution, and chromosomal localization of a novel test
A:Reference number: JC7585; MUID:21092614; PMID:11162526
A:Accession: JC7585
A:Molecule type: mRNA
A:Residues: 1-200 <HUL>
A:Cross-references: GB:AF321127
C:Comment: This protein, a four-transmembrane protein, associates with receptor compl
C:Genetics:
A:Gene: tetm4
A:Map position: 11q12
F:1-48/Domain: cytoplasmic #status predicted <CYT1>
F:49-70/Domain: transmembrane #status predicted <TM1>
F:71-84/Domain: extracellular #status predicted <EXL1>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:106-118/Domain: intracellular #status predicted <INT>
F:119-138/Domain: transmembrane #status predicted <TM3>
F:139-160/Domain: extracellular #status predicted <EXL2>
F:161-182/Domain: transmembrane #status predicted <TM4>
F:183-200/Domain: cytoplasmic #status predicted <CYT2>

Query Match 100.0%; Score 1027; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.5e-87;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVFPPEITA SEYESTELSATSTESTSQPLQKLFARKMKILGTITQILFGIM 60

Db 1 MDSSTAHSPVFLVFPPEITA SEYESTELSATSTESTSQPLQKLFARKMKILGTITQILFGIM 60

QY 61 TFSFGVIFLTLKPPRPFIFLSGYPFWGSLVFNISGAFIAVRKKTETLILSRIM 120

Db 61 TFSFGVIFLTLKPPRPFIFLSGYPFWGSLVFNISGAFIAVRKKTETLILSRIM 120

QY 121 NFISALGAIAIILLTGFILDQNYICGYSHQSOCKAVTVLFGILITLMTFSIELFI 180

Db 121 NFISALGAIAIILLTGFILDQNYICGYSHQSOCKAVTVLFGILITLMTFSIELFI 180

QY 181 SLPFSILGCHSEDCDCQCC 200

Db 181 SLPFSILGCHSEDCDCQCC 200

RESULT 2

```
F;167-183/Disulfide bonds: #status predicted

Query Match          16.5%; Score 169; DB 1; Length 297;
Best Local Similarity 30.3%; Pred. No. 3.4e-08;
Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;

Qy      34 STSQPLOKLFARKMKILGTIOILFGIMTFSGVIFLTLTKPYPRF-PFIFLSCYPEWGS 92
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
```

Db 36 SLVGPQSFMRBRSKTLGAVQIMNGLFHIALGGL-----LMIPAGIYAPICVTWYPLWGG 91

Qy 93 VLFINSAGFLIAVKRKTETLTIILSRIMNFLSALGAIGAILL-----T 136

Db 92 IMYIISGLLAATEKRSKRLVKGMKIMNSLSIFAISGMILSMDILNIKISHFLKMES 151

Qy 137 FGFILQON-YICGY-----SHONS-----QCKAVTVLFGILITLMTFSII-ELFIS 181

Db 152 LNFIRAHPTPIYININCEPANPSEKNSPSTQYCYISQSLFGILSVMLIFAFQELVIA 209

RESULT 5

A31231

IgE Fc receptor beta chain - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999

C:Accession: A31231

R:Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.

Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988

A:Title: Isolation and characterization of cDNAs coding for the beta-subunit of the high

A:Reference number: A31231; MUID:88320465

A:Accession: A31231

A:Molecule type: mRNA

A:Residues: 1-243 <KIN>

A:Cross-references: GB:M22923; GB:J03845; NID:g204116; PIDN:AAA41149.1; PID:g204117

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 15.4%; Score 158; DB 2; Length 243;

Best Local Similarity 28.0%; Pred. No. 2.9e-07;

Matches 45; Conservative 28; Mismatches 72; Indels 16; Gaps 3;

Qy 36 QSPQKLFARKMKILGTIQLFGTTFSGVIFLFTLLKP--YPRFFIFLSGVPFGSV 93

Db 47 QQTWQSFLKKELEFVGTQVLVGLICLCFGFVVCSTIQTDFDEVLVLRAGYPPFGAV 106

Qy 94 LFINSGAFLIAVKRKTETLTIILSRIMNFLSALGAIGAILLTFGFIIDQNYICGYSHQ 153

Db 107 LEVLSGFLSINSRKNTLYLVRSGLGANIVSSIAAGLIAITLNLNSNAYM----- 159

Qy 154 SOCKAVT-----VLFGLITLMTFSIIELFISLPSIL 187

Db 160 NYCKDITEDDGCFTVSTFTELVLMLLFTLILAFCSAVILLII 200

RESULT 6

B34342

IgE Fc receptor beta chain - mouse

C:Species: Mus musculus (house mouse)

C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999

C:Accession: B34342

R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.

J. Biol. Chem. 264, 15323-15327, 1989

A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-Ri) and

A:Reference number: A34342; MUID:89359361

A:Accession: B34342

A:Molecule type: mRNA

A:Status: preliminary

A:Residues: 1-235 <RAC>

A:Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g309225

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 15.3%; Score 157.5; DB 2; Length 235;

Best Local Similarity 30.5%; Pred. No. 3.1e-07;

Matches 60; Conservative 28; Mismatches 82; Indels 27; Gaps 9;

Qy 1 MDSSTAHPVFLVFP-PEITASBYESTELSATFTSQSPLO--KLPARK-MKILGTIQL 56

Db 1 MDTEN-RSRADIALPNQESSAPDIELLEASPAAPKQWRTPLKKELEFLGATQIL 59

Qy 57 FGIWTFSGVIFLTLKPY-----PRFFIFLSGVPFGSVLFINSGAFLIAVKRKTTE 111

Db 60 VGLICLCFGTIVCSVL---YVSDFEDEVLLLYKLGYPFWGAIVLVLSGFLSIISERKNTL 116

Qy 112 TLTIILSRIMNFLSALGAIGAILLTFGFIIDQNYICGYSHONSQCKAVT-----VLFL 164

Db 117 YLVRSGLGANIVSSIAAGTGAILL-----NLTNFAYMNN-CKNVTEDDGCFFVASFT 169

Qy 165 GILITLMTFSIIELFIS 181

Db 170 TELVLMMLFTLILAFCS 186

RESULT 7

A42806

IgE Fc receptor beta chain - human

C:Species: Homo sapiens (man)

C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999

C:Accession: A42806; S21154

R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.

J. Biol. Chem. 267, 12782-12787, 1992

A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta

A:Reference number: A42806; MUID:92316966

A:Accession: A42806

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-244 <KUE>

A:Cross-references: GB:M89796; NID:g337417; PIDN:AAA60269.1; PID:g337418

R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Harada, S.

FEBS Lett. 302, 161-165, 1992

A:Title: Determination of the sequence coding for the beta subunit of the human high-

A:Reference number: S21154; MUID:92339505

A:Accession: S21154

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-244 <MAE>

A:Cross-references: GB:D10583; NID:g219881; PIDN:BAAC01440.1; PID:d1001914; PID:g21988

C:Genetics: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3

A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 12.0%; Score 123.5; DB 2; Length 244;

Best Local Similarity 25.6%; Pred. No. 0.00044;

Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;

Qy 9 PVFLVPPPEITASEYESTELSATFTSQSPLO--KLPARKMKILGTIQLFGIMTFSG 65

Db 21 PAFEVL--EISPOEVSSGRLLKS--ASSPPLHTWLTVLKKEQFLGVTQLTAMICLCFG 76

Qy 66 VIFLFTLLKPYPRPPEI--FLSGYPPFGSVLFINSGAFLIAVKRKTETLTIILSRIMNFL 123

Db 77 TVVCSVLDSIHIEDIFSSFKAGYPPFGAIFFSISGMLSIISERNATYLVRSGLGANTA 136

Qy 124 SALGAIGAILLTFGFIIDQNYICGYSHONS-----QCKAVTVLFGILITLMTF-SIIE 177

Db 137 SSIAGG*GITIL----IINLKKSLAYIHIHSCORFFETKCFMASFSTEIVYVMMFLTILG 192

Qy 178 LFIISLPFSILGCHSE 192

Db 193 LGSVAVSUTICGAGEE 207

RESULT 8

C86844

hypothetical protein ysbC [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: C86844

R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: AB6625; MUID:21235186; PMID:11337471

A:Accession: C86844

A:Status: preliminary

Db 36 SLVGPQSFMRBRSKTLGAVQIMNGLFHIALGGL-----LMIPAGIYAPICVTWYPLWGG 91

Qy 93 VLFINSAGFLIAVKRKTETLTIILSRIMNFLSALGAIGAILL-----T 136

Db 92 IMYIISGLLAATEKRSKRLVKGMKIMNSLSIFAISGMILSMDILNIKISHFLKMES 151

Qy 137 FGFILQON-YICGY-----SHONS-----QCKAVTVLFGILITLMTFSII-ELFIS 181

Db 152 LNFIRAHPTPIYININCEPANPSEKNSPSTQYCYISQSLFGILSVMLIFAFQELVIA 209

RESULT 5

A31231

IgE Fc receptor beta chain - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999

C:Accession: A31231

R:Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.

Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988

A:Title: Isolation and characterization of cDNAs coding for the beta-subunit of the high

A:Reference number: A31231; MUID:88320465

A:Accession: A31231

A:Molecule type: mRNA

A:Residues: 1-243 <KIN>

A:Cross-references: GB:M22923; GB:J03845; NID:g204116; PIDN:AAA41149.1; PID:g204117

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 15.4%; Score 158; DB 2; Length 243;

Best Local Similarity 28.0%; Pred. No. 2.9e-07;

Matches 45; Conservative 28; Mismatches 72; Indels 16; Gaps 3;

Qy 36 QSPQKLFARKMKILGTIQLFGTTFSGVIFLFTLLKP--YPRFFIFLSGVPFGSV 93

Db 47 QQTWQSFLKKELEFVGTQVLVGLICLCFGFVVCSTIQTDFDEVLVLRAGYPPFGAV 106

Qy 94 LFINSGAFLIAVKRKTETLTIILSRIMNFLSALGAIGAILLTFGFIIDQNYICGYSHQ 153

Db 107 LEVLSGFLSINSRKNTLYLVRSGLGANIVSSIAAGLIAITLNLNSNAYM----- 159

Qy 154 SOCKAVT-----VLFGLITLMTFSIIELFISLPSIL 187

Db 160 NYCKDITEDDGCFTVSTFTELVLMLLFTLILAFCSAVILLII 200

RESULT 6

B34342

IgE Fc receptor beta chain - mouse

C:Species: Mus musculus (house mouse)

C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999

C:Accession: B34342

R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.

J. Biol. Chem. 264, 15323-15327, 1989

A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-Ri) and

A:Reference number: A34342; MUID:89359361

A:Accession: B34342

A:Molecule type: mRNA

A:Status: preliminary

A:Residues: 1-235 <RAC>

A:Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g309225

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 15.3%; Score 157.5; DB 2; Length 235;

Best Local Similarity 30.5%; Pred. No. 3.1e-07;

Matches 60; Conservative 28; Mismatches 82; Indels 27; Gaps 9;

Qy 1 MDSSTAHPVFLVFP-PEITASBYESTELSATFTSQSPLO--KLPARK-MKILGTIQL 56

Db 1 MDTEN-RSRADIALPNQESSAPDIELLEASPAAPKQWRTPLKKELEFLGATQIL 59

Qy 57 FGIWTFSGVIFLTLKPY-----PRFFIFLSGVPFGSVLFINSGAFLIAVKRKTTE 111

A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: GB:AE005176; PID:g12724776; PIDN:AAK05853.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ysbC

Query Match 10.3%; Score 106; DB 2; Length 307;
Best Local Similarity 31.9%; Pred. No. 0.023;
Matches 46; Conservative 21; Mismatches 49; Indels 28; Gaps 9;

QY 34 STQSPLOKLPARKKILGTQILFGIMTSEFGVIEL--FTLLKPYPRP--FIFLSGYPPFW 90

DB 16 ANQPIN---ADLRKIVGSPFLASISNF--VGSIFGIITLVTSTLPSPFQFVGSHPAW 71

QY 91 ---GSVLFINSGAFIAVKKRTTETLILSRIMNFSALGAIAIAGIILLTFGLDQNYIC 147

DB 72 IWIGGVLT---GGIFL-----TSNVLLFPR-----LGAVQTVILPILGOILMGTLID 114

QY 148 GYSHONSOCRAVTVL--FLGILITL 170

DB 115 SFGWFHAMQLPMTLMRELGLIITL 138

RESULT 9

Ti6070

hypothetical protein F14B8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16070

R:Geisler, C.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid F14B8.

A:Reference number: Z18456

A:Accession: T16070

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-455 <GEI>

A:Cross-references: EMBL:U028737; NID:g860717; PID:g860719; PIDN:AAA68272.1; CESP:F14B8.3

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F14B8.3

A:Introns: 56/3; 193/2; 224/3; 289/2; 302/3; 370/2; 402/3; 442/3

Query Match

Best Local Similarity 10.0%; Score 103; DB 2; Length 455;

Matches 31; Conservative 15; Mismatches 32; Indels 20; Gaps 3;

QY 49 ILGTQILFGIMT-----FSGVIFLFTLLKPYPR--FPFIFLSGYPPFW 91

DB 303 LLGIVEYFGLISRNHLISICINGISMCAGSVLLPGLQKNPTIYVFLIGY---A 359

QY 92 SVLFINSAGFLIAVKKRTTETLILSRIMNFSALGAI 129

DB 360 SIEFIWDACYLCVVEQVPTVETVGTISGAGSFLSLSGI 397

RESULT 10

JH0751

IgE receptor beta chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: JH0751

R:Bieber, T.; de la Salle, H.; Wollenberg, A.; Hakimi, J.; Chizzonite, R.; Ring, J.; Han

J. Exp. Med. 175, 1285-1290, 1992

A:Title: Human epidermal Langerhans cells express the high affinity receptor for immunog

A:Reference number: JH0751; MUID:92235608

A:Accession: JH0751

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-152 <BIE>

A:Cross-references: GB:X66187; NID:g396463; PIDN:CAA46955.1; PID:g396464

Query Match 9.7%; Score 100; DB 2; Length 152;

Best Local Similarity 27.6%; Pred. No. 0.042; 53; Indels 10; Gaps 3;

Matches 32; Conservative 21; Mismatches 21; Mismatches 53; Indels 10; Gaps 3;

QY 83 FLSGYPPFGSVLFINSAGFLIAVKKRTTETLILSRIMNFSALGAIAIAGIILLTFGLID 142

DB 26 FRAGYPPFGAIFISGMLSIISERRNATVLRGSLGANTASSIAGGTGTTIL---IIN 81

QY 143 QNYICGYSHONS-----QCKRAVTVLFLGILITLMTF--SIIELFISLPSILGCHSE 192

DB 82 LKSLAYIIHHSCKFFETKCFMASFTEIVVMMLFTLILGLSAVSUTICGAGEE 137

RESULT 11

EB1306

probable transmembrane transport protein Cj1040c [imported] - Campylobacter jejuni (s

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000

C:Accession: EB1306

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912

A:Accession: EB1306

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-387 <PAR>

A:Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB73296.1; PID:g696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1040c

Query Match

Best Local Similarity 9.3%; Score 96; DB 2; Length 387;

Matches 42; Conservative 31; Mismatches 65; Indels 30; Gaps 9;

QY 14 PPEITASEYESTELSATTFSTOSQLOKLPARKKILGTIOI--LFGIMTSEFGVIFLFT 71

DB 233 FDKFSTSWLFAQLVAAPVSLFGPL--LLGKURQNLHTFYIAGLCSWVYIAGMLFTFD 290

QY 72 LKPYPPFFI---FLSGYPPFGSVLFINSAGFLIAVKKRTTETLILSRIMNFSALGA 128

DB 291 -----SKITIIISAFIMGFP--WGVF-----GIALLFIAQKSSN-----AQIAARLSALAQ 335

QY 129 TAGIILLTFGLIDQNYICGYSH--ONSCKAVTVLFLGILITLMTF 173

DB 336 GFGYLIAAQG-----QWIIIGLHDKFNFSAIIMLVFVGLVNIIFY 378

RESULT 12

T07892

probable inorganic phosphate transport protein Ptl - barrel medic

C:Species: Medicago truncatula (barrel medic)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001

C:Accession: T07892

R:Liu, H.; Trieu, A.T.; Blaylock, L.A.; Harrison, M.J.

Mol. Plant Microbe Interact. 11, 14-22, 1998

A:Title: Cloning and characterization of two phosphate transporters from Medicago tru

A:Reference number: Z16196; MUID:98086876

A:Accession: T07892

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-537 <LIU>

A:Cross-references: EMBL:AF000354; NID:g2465526; PIDN:AAB81346.1; PID:g2465527

C:Genetics:

A:Gene: Ptl

C:Superfamily: probable inorganic phosphate transport protein PHO84

Search completed: August 27, 2002, 11:04:33
Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2002, 11:02:16 ; Search time 65.04 Seconds
(without alignments)
341.555 Million cell updates/sec

Title: US-09-821-821-2
Perfect score: 1027
Sequence: 1 MDSSTAHSFVFLVFPPEITA.....SLPFLSGHSEDCDCEQCC 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-geneseg/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------------|
| 1 | 1027 | 100.0 | 200 | 22 | AAE13062 Human CD20/IgE-rec |
| 2 | 1027 | 100.0 | 200 | 22 | AAE13062 Human secreted pro |
| 3 | 1004 | 97.8 | 199 | 22 | AAU01210 Novel human membra |
| 4 | 937 | 91.2 | 201 | 22 | AAE10917 Human gene 9 encod |
| 5 | 937 | 91.2 | 212 | 22 | AAE05114 Human reproductive |
| 6 | 735.5 | 71.6 | 149 | 22 | AAU01213 Novel human membra |
| 7 | 687 | 66.9 | 158 | 22 | ABU12234 Human secreted pro |
| 8 | 582 | 56.7 | 138 | 22 | AAU01212 Novel human membra |
| 9 | 414 | 40.3 | 81 | 22 | AAU01211 Novel human membra |
| 10 | 294 | 28.6 | 77 | 21 | AAE03074 Human secreted pro |
| 11 | 263.5 | 25.7 | 67 | 22 | AAE09142 Human secreted pro |

| | | | | | |
|----|-------|------|-----|----|------------------------------|
| 12 | 202.5 | 19.7 | 239 | 20 | AAW96745 High affinity immu |
| 13 | 202.5 | 19.7 | 239 | 21 | AAU50174 Human high affinity |
| 14 | 202.5 | 19.7 | 245 | 21 | AAU94973 Human secreted pro |
| 15 | 201.5 | 19.6 | 220 | 22 | AAU93512 Human polypeptide, |
| 16 | 183.5 | 17.9 | 225 | 21 | AAU94449 Human inflammation |
| 17 | 183.5 | 17.9 | 242 | 21 | AAU73495 Human secreted pro |
| 18 | 183.5 | 17.9 | 248 | 20 | AAU15225 Human receptor pro |
| 19 | 183.5 | 17.9 | 248 | 21 | AAU91531 Human secreted pro |
| 20 | 183.5 | 17.9 | 248 | 22 | AAE12072 Dendritic cell (DC |
| 21 | 183.5 | 17.9 | 248 | 22 | AAE12072 Human hHAIRBs-iso |
| 22 | 183.5 | 17.9 | 250 | 20 | AAU48505 Human breast tumou |
| 23 | 183.5 | 17.9 | 257 | 21 | AAU58419 Lung cancer associ |
| 24 | 183.5 | 17.9 | 273 | 21 | AAU91680 Human secreted pro |
| 25 | 183.5 | 17.9 | 273 | 22 | ABU11989 Human secreted pro |
| 26 | 183.5 | 17.9 | 273 | 22 | AAU25809 Human protein sequ |
| 27 | 183.5 | 17.9 | 273 | 22 | AAU25809 Human colon cancer |
| 28 | 182.5 | 17.8 | 225 | 22 | AAU65272 Human cell surface |
| 29 | 181.5 | 17.7 | 214 | 17 | AAU06503 HTM4 protein. Hom |
| 30 | 181.5 | 17.7 | 214 | 19 | AAU41056 HTM4 protein. Hom |
| 31 | 180 | 17.5 | 226 | 22 | AAU98706 Chandra, a helper |
| 32 | 176.5 | 17.2 | 227 | 22 | AAU19236 Novel human diagno |
| 33 | 172.5 | 16.8 | 267 | 20 | AAU15224 Human receptor pro |
| 34 | 172.5 | 16.8 | 273 | 22 | ABU17004 Novel human diagno |
| 35 | 171 | 16.7 | 178 | 22 | AAE12073 Dendritic cell (DC |
| 36 | 171 | 16.7 | 297 | 13 | AAU20808 Human CD20 antigen |
| 37 | 170.5 | 16.6 | 299 | 21 | AAU91352 Human secreted pro |
| 38 | 169 | 16.5 | 297 | 10 | AAU91356 CD20.4 antigen. |
| 39 | 169 | 16.5 | 297 | 21 | AAU96131 Human lymphocyte c |
| 40 | 169 | 16.5 | 297 | 22 | AAU02440 Human HATERBs isof |
| 41 | 168 | 16.4 | 239 | 22 | AAU74584 Extended human sec |
| 42 | 168 | 16.4 | 248 | 20 | AAU36046 Human HATERBs isof |
| 43 | 168 | 16.4 | 249 | 22 | AAU74583 Human HATERBs isof |
| 44 | 168 | 16.4 | 297 | 17 | AAU91436 Human CD20 antigen |
| 45 | 168 | 16.4 | 297 | 19 | AAU80445 Human CD20.4 antig |

ALIGNMENTS

RESULT 1
AAE13062
ID AAE13062 standard; Protein; 200 AA.
AC AAE13062;
DT 28-JAN-2002 (first entry)
XX Human CD20/IgE-receptor like protein, agp-96614-al.
DE Human; CD20/IgE-receptor like protein; immunoglobulin E; agp-96614-al;
XX agp-96614-al; cancer; abnormal cell proliferation; autoimmune disease;
KW ovarian cancer; brain cancer; arteriosclerosis; vascular restenosis;
KW rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
KW reproductive disease; diabetes; gene therapy.
XX infertility; gene therapy.
XX Homo sapiens.
XX WO200174903-A2.
XX 11-OCT-2001.
XX 29-MAR-2001; 2001WO-US10048.
XX 30-MAR-2000; 2000US-193728P.
XX 27-NOV-2000; 2000US-0723258.
XX (AMGE-) AMGEN INC.
XX Welcher AA, Calzone FJ;
XX WPI: 2001-662968/76.
XX N-PSDB; AAD21441.

```

XX Novel CD20/IgE-receptor like polypeptides and polynucleotides,
PT antagonists and antibodies of the polypeptide useful for treating
PT ameliorating or preventing diseases associated with the polypeptide
PT e.g. cancer, asthma
XX
PS Claim 13; Fig 1; 145pp; English.
XX
CC The invention relates to human CD20/immunoglobulin E (IgE)-receptor
CC like polypeptides designated as agp-96614-al and agp-69406-al and
CC nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are useful for treating, preventing or ameliorating
CC a disease, condition, or disorder which includes cancer such as
CC brain cancer, ovarian cancer; abnormal cell proliferation such as
CC arteriosclerosis, vascular restenosis; pathology from allergens
CC such as allergies, asthma, dermatitis; dysfunction of immune system
CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
CC diabetes, transplant rejection and reproductive diseases such as
CC infertility, preterm labour and delivery, endometriosis etc. They
CC are also useful for identifying antagonists and as immunogens, for
CC raising antibodies which may also be used to prevent, treat or
CC diagnose a number of diseases and disorders. Polynucleotides of the
CC invention are used to map the location of CD20/IgE-receptor like
CC gene and related genes on chromosomes and as hybridisation probes.
CC They are also useful in gene therapy. The present sequence is
CC human CD20/IgE-receptor like protein, agp-96614-al.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 1027; DB 22; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.9e-111;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVFPPEITASEYESTELSATFTSQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 mdsstahspvflvfppeitaseyestelsattfstqspklfarkmkilgtiqlfgim 60

QY 61 TFSFGVIFLTLLKPYRPFIFLPGYFWGSLFINSAGFLIAVKRKTETLILSRIM 120
DB 61 tfsfgvifltllkpyrpfiflpgyfwgslfinsagfliavrkrtetlilsrim 120

QY 121 NPLSALGAIAGIILLTFGLFDQNYICGYSHQNSQCKAVTVLFLGLITLMTFSITELFI 180
DB 121 nplsalgaiagiilltfgldqnyicgyshqnsqckavtvflfgliltmtfsielfi 180

QY 181 SLPPFSILGCHSEDCCEQCC 200
DB 181 slpfsilgchsedcdceqcc 200

RESULT 2
AAG89138
ID AAG89138 standard; Protein; 200 AA.
XX
AC AAG89138;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 258.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
PS 08-DEC-1999; 99US-0169629.
XX

```

```

PR 06-MAR-2000; 2000US-0187470.
XX (GEST ) GENSET.
PA
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
FI WPI; 2001-367870/38.
DR N-PSDB; AAH64741.
DR
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX Claim 21; Page 802-803; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 1027; DB 22; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.9e-111;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVFPPEITASEYESTELSATFTSQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 mdsstahspvflvfppeitaseyestelsattfstqspklfarkmkilgtiqlfgim 60

QY 61 TFSFGVIFLTLLKPYRPFIFLPGYFWGSLFINSAGFLIAVKRKTETLILSRIM 120
DB 61 tfsfgvifltllkpyrpfiflpgyfwgslfinsagfliavrkrtetlilsrim 120

QY 121 NPLSALGAIAGIILLTFGLFDQNYICGYSHQNSQCKAVTVLFLGLITLMTFSITELFI 180
DB 121 nplsalgaiagiilltfgldqnyicgyshqnsqckavtvflfgliltmtfsielfi 180

QY 181 SLPPFSILGCHSEDCCEQCC 200
DB 181 slpfsilgchsedcdceqcc 200

RESULT 3
AAU01210
ID AAU01210 standard; Protein; 199 AA.
XX
AC AAU01210;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #1.
XX
KW Human; membrane protein; membrane receptor; IgE receptor; CD20;
KW Physiological disorder.
XX
OS Homo sapiens.
XX

```


| | | | | | | |
|----------|--|--|-----|----|--------------|-----------------|
| Qy | 61 | TFSRGVIFLFTLLKPYRPPPIFISGYPFWGSVLFINSGAFLIAVKKTKTETLIILSRIM | 120 | PR | 01-SEP-2000; | 2000US-0229344. |
| | | | | PR | 01-SEP-2000; | 2000US-0229345. |
| Db | 61 | tfsrgviflftllkpyrpfifisgypfwgsvlfinsgaflliaavkrkttetliilsrin | 120 | PR | 05-SEP-2000; | 2000US-0229509. |
| | | | | PR | 05-SEP-2000; | 2000US-0229513. |
| Qy | 121 | NFLSALGAIAGIILLTGFILDQNYICGYSHQNSQCKAVTVLFLGILITLMTFSIIELF | 180 | PR | 06-SEP-2000; | 2000US-0230437. |
| | | | | PR | 06-SEP-2000; | 2000US-0230438. |
| Db | 121 | nflsalgalagillltfgildqnyicgyshqnsqckavtrvlfilgtlmtfsiieelfi | 180 | PR | 08-SEP-2000; | 2000US-0231242. |
| | | | | PR | 08-SEP-2000; | 2000US-0231243. |
| Qy | 181 | SLPFSIILG 188 | | PR | 08-SEP-2000; | 2000US-0231244. |
| | | | | PR | 08-SEP-2000; | 2000US-0231413. |
| Db | 181 | slpfsiilgw 188 | | PR | 08-SEP-2000; | 2000US-0231414. |
| | | | | PR | 08-SEP-2000; | 2000US-0232080. |
| RESULT | 5 | | | PR | 08-SEP-2000; | 2000US-0232081. |
| AAM95114 | | | | PR | 12-SEP-2000; | 2000US-0231968. |
| XX | AAM95114 standard; Protein; 212 AA. | | | PR | 14-SEP-2000; | 2000US-0232397. |
| AC | AAM95114; | | | PR | 14-SEP-2000; | 2000US-0232398. |
| DT | | | | PR | 14-SEP-2000; | 2000US-0232399. |
| DE | 21-NOV-2001 (first entry) | | | PR | 14-SEP-2000; | 2000US-0232400. |
| XX | Human reproductive system related antigen SEQ ID NO: 3772. | | | PR | 14-SEP-2000; | 2000US-0232401. |
| XX | Human reproductive system related antigen; reproductive system disorder; | | | PR | 14-SEP-2000; | 2000US-0233063. |
| KW | cancer; gene therapy. | | | PR | 14-SEP-2000; | 2000US-0233064. |
| OS | Homo sapiens. | | | PR | 14-SEP-2000; | 2000US-0233065. |
| XX | | | | PR | 21-SEP-2000; | 2000US-0234223. |
| PN | W0200155320-A2. | | | PR | 21-SEP-2000; | 2000US-0234274. |
| XX | | | | PR | 25-SEP-2000; | 2000US-0234997. |
| PD | 02-AUG-2001. | | | PR | 25-SEP-2000; | 2000US-0234998. |
| XX | | | | PR | 26-SEP-2000; | 2000US-0235484. |
| PF | 17-JAN-2001; 2001WO-US01339. | | | PR | 27-SEP-2000; | 2000US-0235834. |
| XX | | | | PR | 27-SEP-2000; | 2000US-0235836. |
| XX | 31-JAN-2000; 2000US-0179065. | | | PR | 29-SEP-2000; | 2000US-0236327. |
| PR | 04-FEB-2000; 2000US-0180628. | | | PR | 29-SEP-2000; | 2000US-0236367. |
| PR | 24-FEB-2000; 2000US-0184664. | | | PR | 29-SEP-2000; | 2000US-0236368. |
| PR | 02-MAR-2000; 2000US-0186350. | | | PR | 29-SEP-2000; | 2000US-0236369. |
| PR | 16-MAR-2000; 2000US-0189874. | | | PR | 29-SEP-2000; | 2000US-0236370. |
| PR | 17-MAR-2000; 2000US-0190076. | | | PR | 02-OCT-2000; | 2000US-0236802. |
| PR | 18-APR-2000; 2000US-0198123. | | | PR | 02-OCT-2000; | 2000US-0237037. |
| PR | 19-MAY-2000; 2000US-0205515. | | | PR | 02-OCT-2000; | 2000US-0237038. |
| PR | 07-JUN-2000; 2000US-0209467. | | | PR | 02-OCT-2000; | 2000US-0237039. |
| PR | 28-JUN-2000; 2000US-0214886. | | | PR | 02-OCT-2000; | 2000US-0237040. |
| PR | 30-JUN-2000; 2000US-0215135. | | | PR | 13-OCT-2000; | 2000US-0239935. |
| PR | 07-JUL-2000; 2000US-0216647. | | | PR | 13-OCT-2000; | 2000US-0239937. |
| PR | 07-JUL-2000; 2000US-0216880. | | | PR | 20-OCT-2000; | 2000US-0240960. |
| PR | 11-JUL-2000; 2000US-0217487. | | | PR | 20-OCT-2000; | 2000US-0241221. |
| PR | 11-JUL-2000; 2000US-0217496. | | | PR | 20-OCT-2000; | 2000US-0241785. |
| PR | 14-JUL-2000; 2000US-0218290. | | | PR | 20-OCT-2000; | 2000US-0241786. |
| PR | 26-JUL-2000; 2000US-0220963. | | | PR | 20-OCT-2000; | 2000US-0241808. |
| PR | 26-JUL-2000; 2000US-0220964. | | | PR | 20-OCT-2000; | 2000US-0241809. |
| PR | 14-AUG-2000; 2000US-0224518. | | | PR | 20-OCT-2000; | 2000US-0241826. |
| PR | 14-AUG-2000; 2000US-0225213. | | | PR | 01-NOV-2000; | 2000US-0244617. |
| PR | 14-AUG-2000; 2000US-0225214. | | | PR | 08-NOV-2000; | 2000US-0246474. |
| PR | 14-AUG-2000; 2000US-0225266. | | | PR | 08-NOV-2000; | 2000US-0246475. |
| PR | 14-AUG-2000; 2000US-0225267. | | | PR | 08-NOV-2000; | 2000US-0246476. |
| PR | 14-AUG-2000; 2000US-0225268. | | | PR | 08-NOV-2000; | 2000US-0246477. |
| PR | 14-AUG-2000; 2000US-0225270. | | | PR | 08-NOV-2000; | 2000US-0246523. |
| PR | 14-AUG-2000; 2000US-0225447. | | | PR | 08-NOV-2000; | 2000US-0246524. |
| PR | 14-AUG-2000; 2000US-0225757. | | | PR | 08-NOV-2000; | 2000US-0246525. |
| PR | 14-AUG-2000; 2000US-0225758. | | | PR | 08-NOV-2000; | 2000US-0246526. |
| PR | 14-AUG-2000; 2000US-0225759. | | | PR | 08-NOV-2000; | 2000US-0246527. |
| PR | 18-AUG-2000; 2000US-0226279. | | | PR | 08-NOV-2000; | 2000US-0246610. |
| PR | 22-AUG-2000; 2000US-0226681. | | | PR | 08-NOV-2000; | 2000US-0246611. |
| PR | 22-AUG-2000; 2000US-0226868. | | | PR | 08-NOV-2000; | 2000US-0246613. |
| PR | 22-AUG-2000; 2000US-0227182. | | | PR | 17-NOV-2000; | 2000US-0249207. |
| PR | 23-AUG-2000; 2000US-0227009. | | | PR | 17-NOV-2000; | 2000US-0249208. |
| PR | 30-AUG-2000; 2000US-0228924. | | | PR | 17-NOV-2000; | 2000US-0249210. |
| PR | 01-SEP-2000; 2000US-0229287. | | | PR | 17-NOV-2000; | 2000US-0249211. |
| PR | 01-SEP-2000; 2000US-0229343. | | | PR | 17-NOV-2000; | 2000US-0249212. |
| | | | | PR | 17-NOV-2000; | 2000US-0249213. |

PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX N-PSDB; AAL01084.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 11; SEQ ID NO 3772; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention.
XX
XX Sequence 212 AA;
SQ

Query Match 91.2%; Score 937; DB 22; Length 212;
Best Local Similarity 99.5%; Pred. No. 9.4e-101;
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSSTAHPVFLVFPPEITASEYESTELSATTFSTQSPLOKLFARKMKILGTIQLFGIM 60
Db 12 mdsstahspvflvppeitaseyestelsattstqspqlkfarkmkilgtiqlfgim 71

QY 61 TFSFGVIFLTLKPYPRPFIFLSGYFPWGSVLFINSAGFLIAVKRKTETLLILSRIM 120
Db 72 tfsfgviflftlkpyprpfiflsgypfwgsvlfinsgafliavkrktetllilslrim 131

QY 121 NFLSALGATAGIILLTFGLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELFI 180
Db 132 nflsalgalagilllftfgildqnyicgyshonsqckavtvflfgilitmtfsiilelfi 191

QY 181 SLPFSIIIG 188
Db 192 slpfsiwig 199

RESULT 6
AAU01213
ID AAU01213 standard; Protein; 149 AA.
XX
AC AAU01213;
XX

DT 26-SEP-2001 (first entry)
DE Novel human membrane protein #4.
XX
XX Human; membrane protein; membrane receptor; IgE receptor; CD20;
XX Physiological disorder.
XX
XX Homo sapiens.
XX
XX WO200146417-A2.
XX
XX 28-JUN-2001.
XX
XX 12-DEC-2000; 2000WO-US33742.
XX
XX 22-DEC-1999; 99US-0171567.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA;
XX
XX WPI; 2001-408646/43.
XX N-PSDB; AAS04282.
XX
XX Polynucleotide encoding novel human membrane protein, useful for
XX identifying agonist, antagonist or modifiers or for producing
XX antibodies useful in therapeutic, diagnostic and pharmacogenomic
XX applications -
XX
XX Claim 4; Page 31; 32pp; English.
XX
XX The present sequence represents novel human membrane protein #4.
XX Human membrane protein #4 is 1 of 4 human membrane proteins
XX (AAU01210-AAU01213) given in the present invention. These membrane
XX proteins share structural similarity with membrane receptors such as
XX the IgE receptor and mammalian CD20. The novel human membrane proteins
XX are useful for identifying agonists, antagonists and modulators of the
XX membrane proteins. The membrane proteins can be used for diagnosis,
XX drug screening, pharmacogenomic applications, clinical trial monitoring
XX and the treatment of physiological disorders and diseases. The
XX polynucleotides encoding the membrane proteins can be used to generate
XX PCR primers or probes to identify mutations associated with a particular
XX disease.
XX
XX Sequence 149 AA;
SQ

Query Match 71.6%; Score 735.5; DB 22; Length 149;
Best Local Similarity 74.5%; Pred. No. 1.7e-77;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

QY 1 MDSSTAHPVFLVFPPEITASEYESTELSATTFSTQSPLOKLFARKMKILGTIQLFGIM 60
Db 1 mdsstahspvflvppeitaseyestelsattstqspqlkfarkmkilgtiqlfgim 60

QY 61 TFSFGVIFLTLKPYPRPFIFLSGYFPWGSVLFINSAGFLIAVKRKTETLLILSRIM 120
Db 61 tfsfgviflftlkpyprpfiflsgypfwgsvlfinsgafliavkrktet----- 112

QY 121 NFLSALGATAGIILLTFGLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELFI 180
Db 113 -----lgilitmtfsiilelfi 129

QY 181 SLPFSILGHSDDCEQCC 200
Db 130 slpfsilgchsdcdceqcc 149

RESULT 7
ABU12234
ID ABU12234 standard; peptide; 158 AA.
XX

AC ABB12234;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human secreted protein homologue, SEQ ID NO:2604.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer;
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-457740/49.
DR N-PSDB; ABA09478.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20; Page 318; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC

CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC auto-immune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 158 AA;
XX

Query Match 66.9%; Score 687; DB 22; Length 158;
Best Local Similarity 99.3%; Pred. No. 8e-72;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSSTAHSPPVFLVPPPEITASEYESTELSATTFSTOSPLQKLPARKMKILGTITQLIFGIM 60
|||||
Db 12 mdsstahspvflvppeitaseyestelattfstgspkqkfarkmkilgtiqilfglm 71
|||||

QY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRKTTETLILSRIM 120
|||||
Db 72 tfsfgviflftllkpyprpfiflsyypfgsvlfinsgafliavkrktctellilslrim 131
|||||

QY 121 NFLSALGAIAGIILLTFGF 139
|||||
Db 132 nflsalgaiagiilltftf 150
|||||

RESULT 8
AAU01212
ID AAU01212 standard; Protein; 138 AA.
XX
AC AAU01212;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #3.
XX
KW Human; membrane protein; membrane receptor; IgE receptor; CD20;
KW physiological disorder.
XX
OS Homo sapiens.
XX
PN WO200146417-A2.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA;
XX
DR WPI: 2001-408646/43.
DR N-PSDB; AAS04281.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
PS Disclosure; Page 30-31; 32pp; English.
XX
CC The present sequence represents novel human membrane protein #3.
CC Human membrane protein #3 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IgE receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the

CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.

XX Sequence 138 AA;

Query Match 56.7%; Score 582; DB 22; Length 138;
 Best Local Similarity 92.1%; Pred. No. 1.le-59;
 Matches 117; Conservative 4; Mismatches 0; Indels 6; Gaps 1;

QY 1 MDSSTAHSVFLVFPPEITASYESYESTLSATTFTSQSLQKLFARKMKILGTIQLFGIM 60
 |||||
 Db 1 mdsstahspvflvfppeitaseyestelsattftstgslqkifarkmkilgtiqlfgim 60
 |||||
 QY 61 TTSFGVIFLFTLLKPYPRPFIFLSCYFPGSVFLFNSGAPLIAVKKRTTETLI----- 114
 |||||
 Db 61 tfsfgviflftllkpyprpfiflsgypfsgvflfngsapliaavkrkttetlittfvv 120
 |||||
 QY 115 ILSRIMN 121
 ||:::
 Db 121 iltkivs 127

RESULT 9

AAU01211
 ID AAU01211 standard; Protein; 81 AA.

XX AC AAU01211;

XX DT 26-SEP-2001 (first entry)

XX DE Novel human membrane protein #2.

XX KW Human; membrane protein; membrane receptor; IgE receptor; CD20;
 KW physiological disorder.

XX OS Homo sapiens.

XX PN WO200146417-A2.

XX PD 28-JUN-2001.

XX PF 12-DEC-2000; 2000WO-US33742.

XX PR 22-DEC-1999; 99US-0171567.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Turner CA;

XX DR WPI: 2001-408646/43.

XX DR N-PSDB; AAS04280.

XX CC Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -

XX PS Disclosure; Page 30; 32pp; English.

XX CC The present sequence represents novel human membrane protein #2.

XX CC Human membrane protein #2 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IgE receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.

XX Sequence 81 AA;

Query Match 40.3%; Score 414; DB 22; Length 81;
 Best Local Similarity 97.5%; Pred. No. 2e-40;
 Matches 79; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 120 MNFSLAIGALAGIILLTFGFLDQNYICGYSHQNSQCKAVTVLFLGLITLTMFSTIELF 179
 |||||
 Db 1 mnlisalaralagiiilltfgfildqnyicgysghnsgkavtvlfgliltmfsielf 60

QY 180 ISLFFSLGCHSEDCDCQCC 200

Db 61 islpfslgchsedcdceqcc 81

RESULT 10

AAG03074

ID AAG03074 standard; Protein; 77 AA.

XX AC AAG03074;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 7155.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI: 2000-500381/45.

XX DR N-PSDB; AAC03080.

XX CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.

XX CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 77 AA;

Human polypeptide. SEQ ID NO: 3233.

Search completed: August 27, 2002, 11:04:05
Job time: 109 sec

Search completed: August 27, 2002, 11:04:05
Job time: 109 sec

| Result No. | Query 8 | | | ID | Description |
|------------|---------|-------|--------|----|-------------------|
| | Score | Match | Length | | |
| 1 | 86.4 | 11.4 | 1669 | 2 | US-08-910-902A-2 |
| 2 | 86.4 | 11.4 | 1669 | 2 | US-09-213-389-2 |
| 3 | 61.6 | 8.1 | 1661 | 1 | US-08-318-432-3 |
| 4 | 61.6 | 8.1 | 1661 | 1 | US-08-707-340-3 |
| 5 | 61.6 | 8.1 | 1661 | 2 | US-08-994-578-3 |
| 6 | 50.8 | 6.7 | 2545 | 1 | US-07-869-933-22 |
| 7 | 50.8 | 6.7 | 2545 | 4 | US-09-103-663-22 |
| 8 | 44.2 | 5.8 | 7218 | 1 | US-08-232-463-22 |
| 9 | 42.4 | 5.6 | 1738 | 2 | US-08-379-482A-2 |
| 10 | 41 | 5.4 | 658 | 4 | US-09-385-983-327 |
| 11 | 40.4 | 5.3 | 3889 | 2 | US-08-648-298-1 |
| 12 | 40.2 | 5.3 | 1897 | 1 | US-08-184-632-1 |
| 13 | 39.8 | 5.2 | 247 | 1 | US-08-691-641-4 |
| 14 | 39.2 | 5.2 | 2447 | 2 | US-09-014-969-14 |
| 15 | 38.8 | 5.1 | 756 | 2 | US-08-530-165-1 |
| 16 | 38.8 | 5.1 | 1172 | 1 | US-07-945-288-9 |
| 17 | 38.8 | 5.1 | 1172 | 1 | US-08-462-831-9 |
| 18 | 38.8 | 5.1 | 1172 | 1 | US-08-461-809-9 |
| 19 | 38.8 | 5.1 | 1172 | 1 | US-08-461-441-9 |
| 20 | 38.8 | 5.1 | 1172 | 5 | PCS-US93-08518-9 |
| 21 | 38.4 | 5.1 | 991 | 3 | US-08-924-747-25 |
| 22 | 38.4 | 5.1 | 991 | 4 | US-09-247-373B-25 |
| 23 | 38.4 | 5.1 | 991 | 4 | US-09-296-715-25 |
| 24 | 38.4 | 5.1 | 2028 | 3 | US-09-211-930-12 |
| 25 | 38.4 | 5.1 | 2028 | 3 | US-09-340-993-12 |
| 26 | 38.4 | 5.1 | 2028 | 4 | US-09-468-442-12 |
| 27 | 38.4 | 5.1 | 84495 | 4 | US-09-797-906-3 |

11

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
;
US-09-213-389-2

Query Match 11.4%; Score 86.4; DB 2; Length 1669;
Best Local Similarity 51.1%; Pred. No. 3.4e-13;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps

QY 216 aaaaattatttgcagaaaaatgaaaattcttagggactatccagatctctgtttggaatta 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AGAAGTCTTCTGAAGGGAGAACCCAAAGTCCTTGGGGTGTGCAGATTCTGACTGCCCTGA 384

QY 276 tgaacttttcttggaggttatcttctctttcacatttgttaaaccatatccaaagtttc 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 TGAGCGCTTATAGCATGGGAATACCAATGATGTGTATGGCATCTAAATACTTATGGAAGTAAOC 444

QY 336 cctttatattctcttcaggataccactctcgggctcgtgtttgttcaatcaattctcgag 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CTATTTCGGTGTATATCGGTTACACAATTTGGGGTTCAGTAATGTTTATTATTTCAGGAT 504

QY 396 ccttctaatctgactgaaagaaaaaacacagaaactctgataattgagccgaataa 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CTTTGTCAATTTCGACGAGGAATTTAGAACTACAAAGGCCCTGCTCGAGGTAGTCTAGGAA 564

QY 456 tgaattttcttagtgccctggggagcaatagctgggaatcattctcaccatttggttcca 515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 TGAATATACCCAGCTCTGTATCGCTCGATACAGGATCTTAAATCAACACATTTAGCTTGG 624

QY 516 tcttagat-----caaaactacatttgggttatcttcacacaaatagtcaggtgaag 569
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGTTTTATTCAATCCATCACCCCTTACTGTAACTACTATGGAACCTCAATAATTGTCTATG 684

QY 570 ctgttactgtcctgtcttcttgggaattttgattacattgatgaacttcgaactattgaat 629
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGACTATGTCCATCTTAATGGGCTGGATGGCATGGTGCTCTCTTAAGTGTGCTGGAAT 744

QY 630 tattcaattctctgccttctctcaattttggggtg 663

Db 745 TCTGCAATGCTGTGTCCCTCTCTGCTTTGGATG 778

```

TELEPHONE: 415-855-0355

```

3
US-08-318-492-3
: Sequence 3, Application US/08318492
: Patent No. 5352312
: GENERAL INFORMATION:
: APPLICANT: Lim, Bing
: APPLICANT: Adra, Chaker N.
: APPLICANT: Lellias, Jean-Michel
: TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
: TITLE OF INVENTION: ASSAYS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.
: Zip: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,492
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..741
US-08-318-492-3

Query Match 8.1%; Score 61.6; DB 1; Length 1661;
Best Local Similarity 50.6%; Pred. No. 7.1e-07;
Matches 238; Conservative 0; Mismatches 214; Indels 18; Gaps 3;
QY 232 aaaaatgaaatctttagggactccagatcctgttgggaattatgacatttctctttgg 291
DB 234 AAAATTACAAGTCTTGGGGCCATCCAGATCCTGAATGCAGCAATGATTCGGCTTTGGG 293
QY 292 agt-----tatcttcttcttcaacttggtaaaacatatccaaaggttcccttat 342
DB 294 TGCTTTCTGGGTTCCTTGCAATACCCATACCCTTCCAAAAGCACTTCCTTTCTTCAC 353
QY 343 attcttccagatataccattctgggctctgtttgttgcatttaattctggagccttct 402
DB 354 CTTCACAGGCTACCCGATTTGGGGTGTGTTTCTCTCAGTATGAGCAACCTTGTGTC 413
QY 403 aattgcagtgaaagaaacacacagaaactctgataattatggcgcgaataatgaatt 462
DB 414 TGTTGTAGCAGGGATAAACCCACAGCAATGATTCGATGATGATGATGATGATGAT 473
QY 463 tcttagtgcctggagaaatagctgaatcattctctcattctcattcttctcattctc 522
DB 474 TGCCAGTGTACAAATGCGATGGGTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 533
QY 523 tcaaaactacatttgggttat-tctcaccacaaatagtcagtgtaaggtgttact---- 577
DB 534 TATCCAGTCAATTAAGGAGTTGTCACCTTCTCAGAGTCACCGGACCTATGCAATTACAT 593
QY 578 -----gtcctgttcttgggaatttgattacattcagcatttattgaattatt 633
DB 594 GGGCTCCATATCAAAATGCGATGGGTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 653
QY 634 cattctctgcttctcatttggggtgcactcagagattgtgatt 683
DB 654 CGTAATATCTACCATAGCCATGTGTGCAATGCAACATGCTGTAATT 703

RESULT 4
US-08-707-340-3
Sequence 3, Application US/08707340
Patent No. 5705615
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Lelias, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND ASSAYS
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3

Query Match 8.1%; Score 61.6; DB 1; Length 1661;
Best Local Similarity 50.6%; Pred. No. 7.1e-07;
Matches 238; Conservative 0; Mismatches 214; Indels 18; Gaps 3;
QY 232 aaaaatgaaatctttagggactccagatcctgttgggaattatgacatttctctttgg 291
DB 234 AAAATTACAAGTCTTGGGGCCATCCAGATCCTGAATGCAGCAATGATTCGGCTTTGGG 293
QY 292 agt-----tatcttcttcttcaacttggtaaaacatatccaaaggttcccttat 342
DB 294 TGCTTTCTGGGTTCCTTGCAATACCCATACCCTTCCAAAAGCACTTCCTTTCTTCAC 353
QY 343 attcttccagatataccattctgggctctgtttgttgcatttaattctggagccttct 402
DB 354 CTTCACAGGCTACCCGATTTGGGGTGTGTTTCTCTCAGTATGAGCAACCTTGTGTC 413
QY 403 aattgcagtgaaagaaacacacagaaactctgataattatggcgcgaataatgaatt 462
DB 414 TGTTGTAGCAGGGATAAACCCACAGCAATGATTCGATGATGATGATGATGATGATGAT 473
QY 463 tcttagtgcctggagaaatagctgaatcattctctcattctcattcttctcattctc 522
DB 474 TGCCAGTGTACAAATGCGATGGGTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 533
QY 523 tcaaaactacatttgggttat-tctcaccacaaatagtcagtgtaaggtgttact---- 577
DB 534 TATCCAGTCAATTAAGGAGTTGTCACCTTCTCAGAGTCACCGGACCTATGCAATTACAT 593
QY 578 -----gtcctgttcttgggaatttgattacattcagcatttattgaattatt 633
DB 594 GGGCTCCATATCAAAATGCGATGGGTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 653

QY 66 tcactctcttcaattatcacccgacacacatcatgattcagcaccgacacacagtcocgg 125
Db 1094 YY 1153
QY 126 tgttctgtattcttcacgaatacactgcttcagaatagatgacccacagaacttcag 185
Db 1154 YY 1213
QY 186 cacgacactttcaactcaagcccttgcaaaaattatttgcagaaaaatgaaatct 245
Db 1214 YY 1273
QY 246 taggacactatccagatcctgttggaatagacactttcttcttgaggtatctctctt 305
Db 1274 YY 1333
QY 306 tcaactgtttaaaccatcatcaaggttctccctttatattcttcttcaggaatccattct 365
Db 1334 YY 1393
QY 366 ggggctgtttgttcatttaattctgagccttcttaattgcagtgaaaagaaaacca 425
Db 1394 YY 1453
QY 426 cagaactctgataatattgacgaataatgaattttcttagtgccct 474
Db 1454 CTTTAACTACTTGATAGATAGTAATTAACATGATGCTACATGCGCT 1502

RESULT 9
US-08-379-482A-2
; Sequence 2, Application US/08379482A
; Patent No. 5859334
; GENERAL INFORMATION:
; APPLICANT: Brugliera, Filippa
; APPLICANT: Holton, Timothy A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,482A
; FILING DATE: 30-JUL-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9590
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1738 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1413

US-08-379-482A-2

Query Match 5.6%; Score 42.4; DB 2; Length 1738;
Best Local Similarity 57.1%; Pred. No. 0.055;
Matches 97; Conservative 0; Mismatches 71; Indels 2; Gaps 1;
QY 591 gaatttgattacatgatgacttccagcattattgaattattcattctctgcctttct 650
Db 1529 GAATGGCTGTAATAATGCTACTATTGATTGTCAGATAAGGTCACATTTCTACTTCT 1588
QY 651 caatttgggtggtccactcagaggattgtgattgtgaacaattgttctgactagcaactgt 710
Db 1589 CAAGTTTGTGGCACAATCAGCATATGAT--TAAATGAAGATGGTCTTTTACCAGAACATT 1646
QY 711 gagaataaagatgtgttaaaatctcaaaaaaataaaaaaataaaaaa 760
Db 1647 TAAATAAAGATGAGATTCAGTTTAAAAAATAAAAAAATAAAAAA 1696

RESULT 10
US-09-385-982-327/C
; Sequence 327, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 327
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(658)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-327

Query Match 5.4%; Score 41; DB 4; Length 658;
Best Local Similarity 71.6%; Pred. No. 0.092;
Matches 53; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 687 acaatgtgttgactagcactgtgagaataaagatgtgttaaatctcaaaaaa 746
Db 86 ACCTGTCTCTCTGTCATTATTNTGAGATAAATTTCTGTTTCAACCAAAAAA 27
QY 747 aaaaaaataaaaaa 760
Db 26 AAAAAAAAAAAAAA 13

RESULT 11
US-08-648-298-1
; Sequence 1, Application US/08648298
; Patent No. 5871990
; GENERAL INFORMATION:
; APPLICANT: Henrik Clausen
; APPLICANT: Eric Paul Bennett
; TITLE OF INVENTION: UDP-N-acetyl-alpha-D-galactosamine:polypeptide
; TITLE OF INVENTION: N-acetylgalactosaminyltransferase GalNAC-73
; NUMBER OF SEQUENCES: 19


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dardy & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,298
FILING DATE: 15-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4035/08865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212527700
TELEFAX: 2127536237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3889 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Submaxillary gland
US-08-648-298-1

Query Match 5.3%; Score 40.4; DB 2; Length 3889;
Best Local Similarity 48.7%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 116;

QY 529 ctacatttggttattctcaccacaaatagtcagtgtaagcgtgttactgcctgtctt 588
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3653 CTTCTAGTATTTTTTCTTACCAAAATTCACATATCTTGAATGTTGTGATATAAAT 3712
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 ggggaatttgatcattgacgtacttcagcattattgaattcattctctgccttt 648
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3713 TCAATGCGAGATATCTGACTCATTTAAAGCTAAATTTGTACTGATCAATATAAT 3772
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 649 ctcaatttggggtgcactcagagattgtgattgtgaacaaattgttggactagcact 708
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3773 GTAATGGATTTTGACTTTGTAATGATCTCTTTTCATCAAAAGCCTTATTTTCT 3832
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 gtgagaataaagatgtgttaaatctcaaaaaaataaaaaa 754
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3833 AUGTGGAAACACAAATAAAAAATCTCAACACTAAAAA 3878
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-184-632-1
; Sequence 1, Application US/08184632
; Patent No. 5646009
; GENERAL INFORMATION:
; APPLICANT: RHOADS, ROBERT E
; APPLICANT: DEBENEDETTI, ARRICO
; TITLE OF INVENTION: HYBRID VECTOR AND METHOD RESULTING IN
; TITLE OF INVENTION: PROTEIN OVERPRODUCTION BY EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, LeBlanc & Becker
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dardy & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,298
FILING DATE: 15-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4035/08865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212527700
TELEFAX: 2127536237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3889 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Submaxillary gland
US-08-648-298-1

Query Match 5.3%; Score 40.4; DB 2; Length 3889;
Best Local Similarity 48.7%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 529 ctacatttggttattctcaccacaaatagtcagtgtaagcgtgttactgcctgtctt 588
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3653 CTTCTAGTATTTTTTCTTACCAAAATTCACATATCTTGAATGTTGTGATATAAAT 3712
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 ggggaatttgatcattgacgtacttcagcattattgaattcattctctgccttt 648
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3713 TCAATGCGAGATATCTGACTCATTTAAAGCTAAATTTGTACTGATCAATATAAT 3772
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 649 ctcaatttggggtgcactcagagattgtgattgtgaacaaattgttggactagcact 708
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3773 GTAATGGATTTTGACTTTGTAATGATCTCTTTTCATCAAAAGCCTTATTTTCT 3832
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 gtgagaataaagatgtgttaaatctcaaaaaaataaaaaa 754
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3833 AUGTGGAAACACAAATAAAAAATCTCAACACTAAAAA 3878
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-08-691-641-4
; Sequence 4, Application US/08691641
; Patent No. 5728819
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
US-08-184-632-1

Query Match 5.3%; Score 40.2; DB 1; Length 1897;
Best Local Similarity 60.6%; Pred. No. 0.21;
Matches 66; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 652 aatttggggtgcactcagagattgtgattgtgaacaaattgttggactagcactg 711
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1767 AAGTTGGAGCAGTAGTTGATTGTTATTGTTATTGTTATTGTTATTGTTATTGTT 1826
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 712 agataaagatgtgttaaatctcaaaaaaataaaaaa 760
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1827 AGTATCATATATATGGAATAAAAAA 1875
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

US-08-184-632-1

STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,632
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,937
FILING DATE: 05-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: MILLS, DEMETRA J
REGISTRATION NUMBER: 34,506
REFERENCE/DOCKET NUMBER: 434-025B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1897 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-184-632-1
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,641
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-691-641-4

Query Match 5.2%; Score 39.8; DB 1; Length 247;
Best Local Similarity 65.9%; Pred. No. 0.13;
Matches 56; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 676 ttgtgattgtgaacaatgttgttgactgacactgtgagaataaagatgtgttaaatctc 735
Db 163 TTAGGAGTGTATTTTCTACTATGTGTGCGCAGGATACATAAAGANTTTTACAAANTA 222

Qy 736 aaaaaaaaaaaaaaaaaaaaaa 760
Db 223 AAAAAAAAAAAAAAAAAAAAAA 247

RESULT 14

US-09-014-969-14
Sequence 14, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: LaVallie, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 5.2%; Score 39.2; DB 2; Length 2447;
Best Local Similarity 51.1%; Pred. No. 0.4;
Matches 89; Conservative 1; Mismatches 84; Indels 0; Gaps 0;

Qy 587 ttgggaattttgattacattgatcttcagcattattgaattattcattctctgcct 646
Db 2118 TTAGAATGTATAGATAGGTTTGTGAAAGTCTAAATAATGCTGTATAGTATATATAT 2177
Qy 647 ttotcaattttgggggtccactcagagagattgtgtggaacaattgtgtgactagca 706
Db 2178 GGTTCACATATCTGGATCTGTGATTTTGATTTTGTACTTTAAATGTGACAAATAAACCTT 2237
Qy 707 ctgtgagaataaagatgtgttaaatctcaaaaaaaaaaaaaaaaaaaaaa 760
Db 2238 TTGGGAGAAA 2291

RESULT 15

US-08-530-165-1
Sequence 1, Application US/08530165
Patent No. 5907081
GENERAL INFORMATION:
APPLICANT: Isaac, Peter G.
APPLICANT: Roberts, Jeremy A.
APPLICANT: Coupe, Simon A.
TITLE OF INVENTION: Control of Plant Abscission and Pod Dehiscence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: 20005
ZIP: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,165
FILING DATE: 29-SEP-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/00689
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9306726.2
FILING DATE: 31-MAR-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 756 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 15..458
US-08-530-165-1

| | Query Match | 5.1% | Score 38.8; | DB 2; | Length 756; |
|----|-----------------------|--|-------------|------------|-------------|
| | Best Local Similarity | 56.2%; | Pred. No. | 0.35; | |
| | Matches | 73; Conservative | 0; | Mismatches | 57; Indels |
| | | | | Gaps | 0; |
| Qy | 631 | atcattcttcgaccttccaattttgggggtgccactcgagagtgtgttgtagca | 690 | | |
| | | | | | |
| Dd | 618 | ATTGTTTTTATTTGTTCCTTAATTATAATACATGCACAGGGTGTTGTTACTGG | 677 | | |
| | | | | | |
| Qy | 691 | atgttgttgactagcactgagaataaagtgtgttaaatctcaaaaaaa | 750 | | |
| | | | | | |
| Dd | 678 | TGTGTGTTAAATTAATAAATAATCTTAACTCAAAAAAAA | 737 | | |
| | | | | | |
| Qy | 751 | aaaaaaaaa | 760 | | |
| | | | | | |
| Dd | 738 | AAAAAAAAA | 747 | | |
| | | | | | |

Search completed: August 30, 2002, 02:42:19
Job time: 7102 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2002, 01:33:42 ; Search time 214.01 Seconds
(without alignments)
6097.168 Million cell updates/sec

Title: US-09-821-821-1

Perfect score: 760

Sequence: 1 ttccagctcctccagcagcc.....aaaaaaaaaaaaaaaaaaaaa 760

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------------|
| 1 | 760 | 100.0 | 760 | 22 | AAD21441 Human CD20/IgE-rec |
| 2 | 730.2 | 96.1 | 747 | 22 | AAH64741 Human secreted pro |
| 3 | 689.8 | 90.8 | 737 | 22 | AAL01084 Human reproductive |
| 4 | 668.4 | 87.9 | 689 | 22 | AAL18275 Human immune syste |
| 5 | 666.4 | 87.7 | 676 | 22 | AAS04283 DNA sequence encod |
| 6 | 635.8 | 83.7 | 695 | 22 | ABA09478 Human secreted pro |
| 7 | 598.2 | 78.7 | 603 | 22 | AAS04279 Novel human membra |
| 8 | 351.2 | 46.2 | 372 | 22 | AAC03080 Human secreted pro |
| 9 | 347.4 | 45.7 | 450 | 22 | AAS04282 Novel human membra |

| | | | | | |
|----|-------|------|------|----|-----------------------------|
| 10 | 340.2 | 44.8 | 417 | 22 | AAS04281 Novel human membra |
| 11 | 242.8 | 31.9 | 246 | 22 | AAS04280 Novel human membra |
| 12 | 233.4 | 30.7 | 468 | 22 | AAH64745 Human secreted pro |
| 13 | 86.4 | 11.4 | 1330 | 21 | AAI16693 Human secreted pro |
| 14 | 86.4 | 11.4 | 1669 | 20 | AAI15104 High affinity immu |
| 15 | 86.4 | 11.4 | 1670 | 21 | AAZ32842 Human high affinit |
| 16 | 84.8 | 11.2 | 1677 | 22 | AAK94442 Human full-length |
| 17 | 78.8 | 10.4 | 822 | 22 | AAK92131 Human cDNA 5'-end |
| 18 | 78.8 | 10.4 | 822 | 22 | AAK93395 Human IGERB homolo |
| 19 | 69.4 | 9.1 | 562 | 22 | ABA09445 Human IGERB homolo |
| 20 | 64.4 | 8.5 | 1710 | 23 | AAH83423 DNA encoding a bo |
| 21 | 62.6 | 8.2 | 1563 | 21 | AAZ36238 Murine wild-type F |
| 22 | 61.8 | 8.1 | 708 | 22 | AAH77694 HTM4 protein codin |
| 23 | 61.6 | 8.1 | 1661 | 17 | AAH45120 HTM4 gene. Homo s |
| 24 | 61.6 | 8.1 | 1661 | 19 | AAV03875 Clone CD20.4 encod |
| 25 | 59.8 | 7.9 | 1474 | 13 | AAO21170 Human CD20.4 ant |
| 26 | 58.2 | 7.7 | 1473 | 10 | AAH90610 Human CD20.4 ant |
| 27 | 58.2 | 7.7 | 1474 | 17 | AAH14710 Human CD20.4 ant |
| 28 | 58.2 | 7.7 | 1474 | 19 | AAV63448 Human CD20.4 ant |
| 29 | 58.2 | 7.7 | 1474 | 20 | AAV81206 Human CD20.4 ant |
| 30 | 58.2 | 7.7 | 1474 | 21 | AAH05084 Human cell surface |
| 31 | 58.2 | 7.7 | 1474 | 22 | AAH03178 Human lymphocyte c |
| 32 | 58.2 | 7.7 | 1476 | 10 | AAH90613 CD20 cDNA. AAN90 |
| 33 | 58.2 | 7.7 | 1476 | 13 | AAO21173 Human CD20 antigen |
| 34 | 58.2 | 7.7 | 1476 | 17 | AAH14713 Human CD20 antigen |
| 35 | 58.2 | 7.7 | 1476 | 19 | AAV63451 Human CD20 antigen |
| 36 | 58.2 | 7.7 | 1476 | 20 | AAV81209 Human CD20 antigen |
| 37 | 58.2 | 7.7 | 1476 | 21 | AAH05087 Human cell surface |
| 38 | 58.2 | 7.7 | 1476 | 22 | AAH03181 Human lymphocyte c |
| 39 | 54.6 | 7.2 | 1060 | 20 | AAH97730 Extended human sec |
| 40 | 54.2 | 7.1 | 932 | 22 | AAH19723 Dendritic cell (DC |
| 41 | 54.2 | 7.1 | 945 | 19 | AAV59668 Human secreted pro |
| 42 | 54.2 | 7.1 | 1001 | 21 | AAZ52580 Human secreted pro |
| 43 | 54.2 | 7.1 | 1036 | 22 | AAH19722 Dendritic cell (DC |
| 44 | 54.2 | 7.1 | 1106 | 22 | ABA09233 Human secreted pro |
| 45 | 54.2 | 7.1 | 1106 | 22 | AAH93750 Human protein enco |

ALIGNMENTS

RESULT 1

AAD21441
ID AAD21441 standard; cDNA; 760 BP.

XX

AC AAD21441;

XX

XX 28-JAN-2002 (first entry)

DT

XX

DE Human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.

XX

Human; CD20/IgE-receptor like protein; immunoglobulin E; agp-96614-al;
agp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;
ovarian cancer; brain cancer; arteriosclerosis; vascular stenosis;
rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
reproductive disease; diabetes; transplant rejection; endometriosis;
infertility; gene therapy; ss.

XX Homo sapiens.

OS

XX

PH Key

FT CDS

FT Location/Qualifiers

FT 98..700

FT /tag= a

FT /product= "Human CD20/IgE-receptor like protein,

FT agp-96614-al"

FT

XX W0200174903-A2.

XX

XX 11-OCT-2001.

XX

XX 29-MAR-2001; 2001WO-US10048.

XX

XX 30-MAR-2000; 2000US-193728P.

PR

PR 27-NOV-2000; 2000US-0723258.
PA (AMGE-) AMGEN INC.
XX Welcher AA, Calzone FJ;
XX
XX
XX WPI: 2001-662968/76.
XX P-PSDB; AAE13062.
XX
XX Novel CD20/IgE-receptor like polypeptides and polynucleotides,
XX antagonists and antibodies of the polypeptide useful for treating
XX ameliorating or preventing diseases associated with the polypeptide
XX e.g. cancer, asthma
XX
XX Claim 1; Fig 1; 145pp; English.
XX
XX The invention relates to human CD20/immunoglobulin E (IgE)-receptor
XX like polypeptides designated as agp-96614-al and agp-69406-al and
XX nucleic acid molecules encoding such polypeptides. Polypeptides of
XX the invention are useful for treating, preventing or ameliorating
XX a disease, condition, or disorder which includes cancer such as
XX brain cancer, ovarian cancer; abnormal cell proliferation such as
XX arteriosclerosis, vascular restenosis; pathology from allergens
XX such as allergies, asthma, dermatitis; dysfunction of immune system
XX such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
XX diabetes, transplant rejection and reproductive diseases such as
XX infertility, preterm labour and delivery, endometriosis etc. They
XX are also useful for identifying antagonists and as immunogens, for
XX raising antibodies which may also be used to prevent, treat or
XX diagnose a number of diseases and disorders. Polynucleotides of the
XX invention are used to map the location of CD20/IgE-receptor like
XX gene and related genes on chromosomes and as hybridisation probes.
XX They are also useful in gene therapy. The present sequence is
XX human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.
XX
SQ Sequence 760 BP; 232 A; 157 C; 126 G; 245 T; 0 other;

Query Match 100.0%; Score 760; DB 22; Length 760;
Best Local Similarity 100.0%; Pred. No. 5.2e-165;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttccagtgctccaggcagctcagcacagaagaacatggtctagactgaagtaccaa 60
Db 1 ttccagtgctccaggcagctcagcacagaagaacatggtctagactgaagtaccaa 60
Qy 61 cttaatcatctctcttcaaatattacccgacacatcatggtattcagccgcacacag 120
Db 61 cttaatcatctctcttcaaatattacccgacacatcatggtattcagccgcacacag 120
Qy 121 tccgggtttctgttatttctccagaaatcaactgttccagaatatgagtcacagaact 180
Db 121 tccgggtttctgttatttctccagaaatcaactgttccagaatatgagtcacagaact 180
Qy 181 ttccagccagacgttttcaactcaaacgcttgcgaaataattatttgcagaaatgaa 240
Db 181 ttccagccagacgttttcaactcaaacgcttgcgaaataattatttgcagaaatgaa 240
Qy 241 aattctggagactatccagatcctgtttggaattatgaccttttcttttggagtattctt 300
Db 241 aattctggagactatccagatcctgtttggaattatgaccttttcttttggagtattctt 300
Qy 301 ccttttcaactttgtaaaacatccaaggtttcccttttatatttcttccagatatcc 360
Db 301 ccttttcaactttgtaaaacatccaaggtttcccttttatatttcttccagatatcc 360
Qy 361 attctggggctgtttgttctattattctggagccttctactaatgtcagtgaagaaga 420
Db 361 attctggggctgtttgttctattattctggagccttctactaatgtcagtgaagaaga 420
Qy 421 aaccacagaactctgataatattgagccggaataatgaattttcttagtgccttggagc 480
Db 421 aaccacagaactctgataatattgagccggaataatgaattttcttagtgccttggagc 480

QY 481 aatagctgggaatcattctcctcacatttggtttccatcctagatcaaaactacatttgg 540
Db 481 aatagctgggaatcattctcctcacatttggtttccatcctagatcaaaactacatttgg 540
QY 541 ttattctcaccacaaatagtcagtgtaaggctgttactgtcttcttcttgggaatttggat 600
Db 541 ttattctcaccacaaatagtcagtgtaaggctgttactgtcttcttcttgggaatttggat 600
QY 601 tacatgtagactttcagcaattattgaattattcattcttcttcccttcttcaatttgg 660
Db 601 tacatgtagactttcagcaattattgaattattcattcttcttcccttcttcaatttgg 660
QY 661 gtgccactcagagattgtgattgtgaacaatgttctgactagcactgtgagaataag 720
Db 661 gtgccactcagagattgtgattgtgaacaatgttctgactagcactgtgagaataag 720
QY 721 atgtgttaaaatctcaaaaaaataaaaaaataaaaaa 760
Db 721 atgtgttaaaatctcaaaaaaataaaaaaataaaaaa 760
RESULT 2
AAH64741
ID AAH64741 standard; cDNA; 747 BP.
XX
AC AAH64741;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 17.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
XX Homo sapiens.
XX
XX WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
XX 08-DEC-1999; 99US-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI: 2001-367870/38.
DR P-PSDB; AAG89138.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
XX proteins, useful in gene therapy and vaccination against a variety of
XX diseases, and for diagnosis of those diseases -
PS Claim 7; Page 576-577; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased
XX GENSET gene expression by rectifying mutations or deletions in a
XX patient's genome that affect the activity of GENSET or by supplementing
XX the patients own production of GENSET polypeptides. Conversely,
XX antisense nucleic acid molecules may be administered to down regulate
XX GENSET expression by binding with the cells' own genes and preventing
XX their expression. The sense and antisense nucleic acids may also be
XX used as DNA probes in diagnostic assays to detect and quantitate the
XX presence of similar nucleic acid sequences in samples, and hence to
XX determine which patients may be in need of restorative therapy.

CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX
SQ Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;

| | | | | | |
|---------------------------|-----|---|---------------------|--------|-------------|
| Query Match | | 96.1%; | Score 730.2; | DB 22; | Length 747; |
| Best Local Similarity | | 99.6%; | Pred. No. 3.5e-158; | | |
| Matches 732; Conservative | | 0; | Mismatches 3; | Indels | 0; Gaps 0; |
| QY | 26 | acaagaaaaaacatggtctagactgaagtacacaaataaataatcctcttcacaaattatc | 85 | | |
| Db | 1 | acaagaaaaaacatggtctagactgaagtacacaaataaataatcctcttcacaaattatc | 60 | | |
| QY | 86 | accgacaccatcatgattcagaccgacacagtcggtgttctggtatttctcca | 145 | | |
| Db | 61 | accgacaccatcatgattcagaccgacacagtcggtgttctggtatttctcca | 120 | | |
| QY | 146 | gaatcactgcttcagaatatagtcacagacacttccagccacacacacttttcaactcaa | 205 | | |
| Db | 121 | gaatcactgcttcagaatatagtcacagacacttccagccacacacacttttcaactcaa | 180 | | |
| QY | 206 | agcccttgcaaaaattattctagaaaaatgaaaaatcttagggactatccagatcctg | 265 | | |
| Db | 181 | agcccttgcaaaaattattctagaaaaatgaaaaatcttagggactatccagatcctg | 240 | | |
| QY | 266 | tttggaaatgatgaccttttttggagtattcttcttccactttgttaaaacccatat | 325 | | |
| Db | 241 | tttggaaatgatgaccttttttggagtattcttcttccactttgttaaaacccatat | 300 | | |
| QY | 326 | ccaaggttccctttatattcttcaggatatccattctgggctctgtttgttgcatt | 385 | | |
| Db | 301 | ccaaggttccctttatattcttcaggatatccattctgggctctgtttgttgcatt | 360 | | |
| QY | 386 | aattctggagccttccctaaattgcagtgaagaaaaaacacagaaaactctgataatgtg | 445 | | |
| Db | 361 | aattctggagccttccctaaattgcagtgaagaaaaaacacagaaaactctgataatgtg | 420 | | |
| QY | 446 | agccgaataatgaatttcttagtgccttgggagcaaatagctgggaatcttctctccaca | 505 | | |
| Db | 421 | agccgaataatgaatttcttagtgccttgggagcaaatagctgggaatcttctctccaca | 480 | | |
| QY | 506 | tttggtttccctcagataaaactaatgttggtttattctccacaaaatagtcagtgt | 565 | | |
| Db | 481 | tttggtttccctcagataaaactaatgttggtttattctccacaaaatagtcagtgt | 540 | | |
| QY | 566 | aaggctgttactgctcctgttcttgggaattttgattacattgatgactttcagcattatt | 625 | | |
| Db | 541 | aaggctgttactgctcctgttcttgggaattttgattacattgatgactttcagcattatt | 600 | | |
| QY | 626 | gaattatcattctctgcctttctcattttgggggtgcacctcagaggattgtgattgt | 685 | | |
| Db | 601 | gaattatcattctctgcctttctcattttgggggtgcacctcagaggattgtgattgt | 660 | | |
| QY | 686 | gaacatgtttgtactagactgtgagataaagatgtgttaaaatctcaaaaaaaaa | 745 | | |
| Db | 661 | gaacatgtttgtactagactgtgagataaagatgtgttaaaatctcaaaaaaaaa | 720 | | |
| QY | 746 | aaaaaaaaaaaaaa 760 | | | |
| Db | 721 | aaaaaaaaaaaaaa 735 | | | |

RESULT 3
AAL01084/c
ID AAL01084 standard; cDNA; 737 BP.
XX
AC AAL01084;
XX
DT 21-NOV-2001 (first entry)
XX

| | |
|----|---|
| DE | Human reproductive system related antigen cDNA SEQ ID NO: 1085. |
| XX | |
| KW | Human; reproductive system related antigen; reproductive system disorder; |
| KW | cancer; gene therapy; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200155320-A2. |
| XX | |
| PD | 02-AUG-2001. |
| XX | |
| PF | 17-JAN-2001; 2001WO-US01339. |
| XX | |
| PR | 31-JAN-2000; 2000US-0179065. |
| PR | 04-FEB-2000; 2000US-0180628. |
| PR | 24-FEB-2000; 2000US-0184664. |
| PR | 02-MAR-2000; 2000US-0186350. |
| PR | 16-MAR-2000; 2000US-0189874. |
| PR | 17-MAR-2000; 2000US-0190076. |
| PR | 18-APR-2000; 2000US-0198123. |
| PR | 19-MAY-2000; 2000US-0205515. |
| PR | 07-JUN-2000; 2000US-0209467. |
| PR | 28-JUN-2000; 2000US-0214886. |
| PR | 30-JUN-2000; 2000US-0215135. |
| PR | 07-JUL-2000; 2000US-0216647. |
| PR | 07-JUL-2000; 2000US-0216880. |
| PR | 11-JUL-2000; 2000US-0217487. |
| PR | 11-JUL-2000; 2000US-0217496. |
| PR | 14-JUL-2000; 2000US-0218290. |
| PR | 26-JUL-2000; 2000US-0220963. |
| PR | 26-JUL-2000; 2000US-0220964. |
| PR | 14-AUG-2000; 2000US-0224518. |
| PR | 14-AUG-2000; 2000US-0224519. |
| PR | 14-AUG-2000; 2000US-0225213. |
| PR | 14-AUG-2000; 2000US-0225214. |
| PR | 14-AUG-2000; 2000US-0225266. |
| PR | 14-AUG-2000; 2000US-0225267. |
| PR | 14-AUG-2000; 2000US-0225268. |
| PR | 14-AUG-2000; 2000US-0225270. |
| PR | 14-AUG-2000; 2000US-0225447. |
| PR | 14-AUG-2000; 2000US-0225757. |
| PR | 14-AUG-2000; 2000US-0225758. |
| PR | 14-AUG-2000; 2000US-0225759. |
| PR | 18-AUG-2000; 2000US-0226279. |
| PR | 22-AUG-2000; 2000US-0226681. |
| PR | 22-AUG-2000; 2000US-0226688. |
| PR | 22-AUG-2000; 2000US-0227182. |
| PR | 23-AUG-2000; 2000US-0227009. |
| PR | 30-AUG-2000; 2000US-0228924. |
| PR | 01-SEP-2000; 2000US-0229287. |
| PR | 01-SEP-2000; 2000US-0229343. |
| PR | 01-SEP-2000; 2000US-0229344. |
| PR | 01-SEP-2000; 2000US-0229345. |
| PR | 05-SEP-2000; 2000US-0229509. |
| PR | 05-SEP-2000; 2000US-0229513. |
| PR | 06-SEP-2000; 2000US-0230437. |
| PR | 06-SEP-2000; 2000US-0230438. |
| PR | 08-SEP-2000; 2000US-0231242. |
| PR | 08-SEP-2000; 2000US-0231243. |
| PR | 08-SEP-2000; 2000US-0231244. |
| PR | 08-SEP-2000; 2000US-0231413. |
| PR | 08-SEP-2000; 2000US-0231414. |
| PR | 08-SEP-2000; 2000US-0232080. |
| PR | 08-SEP-2000; 2000US-0232081. |
| PR | 12-SEP-2000; 2000US-0231968. |
| PR | 14-SEP-2000; 2000US-0232397. |
| PR | 14-SEP-2000; 2000US-0232398. |
| PR | 14-SEP-2000; 2000US-0232399. |
| PR | 14-SEP-2000; 2000US-0232400. |
| PR | 14-SEP-2000; 2000US-0232401. |
| PR | 14-SEP-2000; 2000US-0233063. |
| PR | 14-SEP-2000; 2000US-0233064. |
| PR | 14-SEP-2000; 2000US-0233065. |

PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-024475.
PR 08-NOV-2000; 2000US-024476.
PR 08-NOV-2000; 2000US-024477.
PR 08-NOV-2000; 2000US-024478.
PR 08-NOV-2000; 2000US-024523.
PR 08-NOV-2000; 2000US-024524.
PR 08-NOV-2000; 2000US-024525.
PR 08-NOV-2000; 2000US-024526.
PR 08-NOV-2000; 2000US-024527.
PR 08-NOV-2000; 2000US-024528.
PR 08-NOV-2000; 2000US-024532.
PR 08-NOV-2000; 2000US-024533.
PR 08-NOV-2000; 2000US-024609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
DR P-PSDB; AAM95114.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition
XX
XX
PS Claim 1; SEQ ID NO 1085; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 737 BP; 240 A; 128 C; 152 G; 217 T; 0 other;

Query Match 90.8%; Score 689.8; DB 22; Length 737;
Best Local Similarity 98.9%; Pred. No. 6.4e-149;
Matches 705; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 47 gactgaagtagcacaataatcatctctcttcaataatataccgcacacatcatgattca 106
Db 737 GACTGAAGTAGCACAATAATCATCTCTCTCTTCAAAATATATCAGCAGACCATGATTC 678
Qy 107 agcaacgcacacagtcgcggtgtttctgtattctccagaaatcactgcttcagaatat 166
Db 677 AGCACCACACAGTCGCGGTGTTCGTGATTTCTCCAGAAATCAGTCTCAGATAT 618
Qy 167 gactccagagaactttcagccagacacttttcaactcaaaagcccttgcaaaaattttt 226
Db 617 GAGTCCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCACAAATATTT 558
Qy 227 gctagaaaaatgaaatccttaggagctatccagatccctgttttggaattatgacctttct 286
Db 557 GCTAGAAAAATGAAATCTTAGGAGCTATCCAGATCCCTGTTGGAAATATGACCTTTCT 498
Qy 287 ttggagttatctctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 346
Db 497 TTTGGAGTTATCTCTCTCTTTCACCTTGTGTTAAACCATATCCAAAGCTTTCCCTTTATATTT 438
Qy 347 ctttcaggatattcattctctgggctctgtttgttcttcttcttcttcttcttcttcttct 406
Db 437 CTTTCAGGATATCCATTCIGGGGCTCTGTTTGTTCATTAATTCGAGGCTTCTCTAATT 378
Qy 407 gcagtgaaagaaaaaacccagagaactctgataataattgagccgaataatgaattttctt 466
Db 377 GCAGTGAAAGAAAAAACCCAGAAACTCTGATTAATTTAGCCCGCAATAATGAATTTCTTT 318
Qy 467 agtgcctctggagcaatagctggaatcattctcctcacatttggtttcttctctagatcaa 526
Db 317 AGTGCCTGGAGCAATAGCTGGAATCATCTCTCCACATTTGGTTTCATCCTTAGATCAA 258
Qy 527 aactacatttgggtatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 586
Db 257 AACTACATTTTGGTTATTCTTCCAAAAATAGTCAGTGTAAAGGCTGTACTGTCTCTGTTTC 198
Qy 587 ttgggaatttggattacattgagctttcagcattattgaattattcattctctctgctt 646
Db 197 TTGGGAATTTTGATTTACATTTGAGCTTTTCAGCATTTATGAATTTATTTCTCTGCTT 138
Qy 647 ttctcaatttgggggtgcgaactcagaggattgtgattgtgaacaatgtgttgactagcaa 706
Db 137 TTCTCAA-TTTGGGGTGCCACTCAGAGGATTGTGATTGTGACAAATGTTGTGACTAGCA 79

QY 707 ctgtgagataaagatgtgttaaaatctcaaaaaaataaaaaaataaaaaa 759
 Db 78 CTGTGAGATAAAGATGTGTAAATATTATGTATAAAAAAATAAAAAA 26

RESULT 4

AA018275
 ID AAD18275 standard; cDNA; 689 BP.

XX AC AAD18275;

XX DT 18-DEC-2001 (first entry)

XX DE Human immune system-related protein-encoding gene 9 cDNA clone HTEN45.

XX KW Human; immune system-related protein; allergy; rheumatoid arthritis;
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cystostatic;
 KW diabetes mellitus; arrhythmia; wound healing; ischaemic lesion; AIDS;
 KW Acquired Immune Deficiency Syndrome; virucide; hepatotropic; vasotropic;
 KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
 KW wound healing; cell proliferation; skin aging; endocrine disorder;
 KW food preservative; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 29..634

XX FT /*tag= a

XX FT /product= "Human immune-system related protein"

XX WO200166722-A1.

XX 13-SEP-2001.

XX 07-MAR-2001; 2001WO-US07260.

XX 08-MAR-2000; 2000US-187873P.

XX 11-AUG-2000; 2000US-224367P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX N1 J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;

XX Gruber JR, Endress GA, Ruben SM;

XX WPT: 2001-589939/66.

XX P-PSDB: AAE10917.

XX Novel isolated immune system-related polypeptide useful for treating
 XX rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
 XX diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
 XX viral hepatitis

XX Claim 1; Page 301; 315pp; English.

XX The invention relates to human immune system-related protein and their
 CC DNA. Human immune-system related protein and DNA are useful for
 CC preventing, treating or ameliorating a medical condition in a mammalian
 CC subject, for diagnosing, preventing or treating immune system-associated
 CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
 CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
 CC (allergies), infectious diseases (e.g., viral hepatitis), complement
 CC activation disorders, immune complex diseases, neoplastic disorders
 CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
 CC associated with neovascularisation, diseases at the cellular level,
 CC cardiovascular disorders (arrhythmias), wound healing and epithelial
 CC cell proliferation, endocrine disorders (diabetes mellitus) and
 CC neurological disorders (ischemic lesions). Immune-system related protein
 CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
 CC modulate mammalian characteristics, to change a mammal's mental or
 CC physical state, or as a food additive or preservative. Immune-system
 CC related DNA is useful in gene therapy, for chromosome identification,
 CC radiation hybrid mapping, long range restriction mapping and in forensic

CC bioloq. The present sequence represents a human immune-system related
 CC protein-encoding cDNA of the invention.
 XX
 SQ Sequence 689 BP; 204 A; 138 C; 114 G; 233 T; 0 other;

Query Match 87.9%; Score 668.4; DB 22; Length 689;

Best Local Similarity 99.0%; Pred. No. 5.1e-144;

Matches 683; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 70 ctctttcaaatattatcacccagacacacatcatgattcgaagccacacacatcgggtgtt 129

Db 1 ctctttcaaatattatcacccagacacacatcatgattcgaagccacacacatcgggtgtt 60

QY 130 ctctgtattctccagaaatcaactgcttcagaatattgctccagaaatcttcagccac 189

Db 61 ttctgtattctccagaaatcaactgcttcagaatattgctccagaaatcttcagccac 120

QY 190 gacctttcaactcaaaagcccttcgaaaaattatttctgtagaaaaatgaaaaattcttagg 249

Db 121 gacctttcaactcaaaagcccttcgaaaaattatttctgtagaaaaatgaaaaattcttagg 180

QY 250 gactatccagatcctgtttggaattatgacctttctttctgttgattctctctttccac 309

Db 181 gactatccagatcctgtttggaattatgacctttctttctgttgattctctctttccac 240

QY 310 ttctgttaaaacacatataccaaaggtttccctttatatattcttcagatatacattctgtggg 369

Db 241 ctgtgttaaaacacatataccaaaggtttccctttatatattcttcagatatacattctgtggg 300

QY 370 ctctgtttgttctcaataattctgtagccttctctaaatttcagtgaaaaaaccacaga 429

Db 301 ctctgtttgttctcaataattctgtagccttctctaaatttcagtgaaaaaaccacaga 360

QY 430 aactctgataattatgagccgaataataatttcttctgtagccttgagcaatagctgg 489

Db 361 aactctgataattatgagccgaataataatttcttctgtagccttgagcaatagctgg 420

QY 490 aatcattctctccacatttgggtttctcatctcagatatacaaaactacatttgggttattctca 549

Db 421 aatcattctctccacatttgggtttctcatctcagatatacaaaactacatttgggttattctca 480

QY 550 caaaaatagtcagtgtaaggctgttactctctgttcttcttgggaatttggattacattgat 609

Db 481 caaaaatagtcagtgtaaggctgttactctctgttcttcttgggaatttggattacattgat 540

QY 610 gactttcagcattattgaattattctctctgctttctcaattttgggtgcaactc 669

Db 541 gactttcagcattattgaattattctctctgctttctcaattttgggtgcaactc 599

QY 670 agaggattgtgattgtgaacaattgttctgactagcactgtgagaataaagaatgtgttaa 729

Db 600 agaggattgtgattgtgaacaattgttctgactagcactgtgagaataaagaatgtgttaa 659

QY 730 aatctcaaaaaaataaaaaaataaaaaa 759

Db 660 aatattatgtaaaaaaataaaaaaataaaaaa 689

RESULT 5

AA004283

ID AAS04283 standard; DNA; 676 BP.

XX AC AAS04283;

XX DT 26-SEP-2001 (first entry)

XX DE DNA sequence encoding novel human membrane protein.

XX KW Human; membrane protein; membrane receptor; IgE receptor; CD20;

XX KW physiological disorder; ds.

XX OS Homo sapiens.

CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis
CC depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ

Sequence 695 BP; 203 A; 142 C; 116 G; 234 T; 0 other;

Query Match 83.78; Score 635.8; DB 22; Length 695;
Best Local Similarity 98.4%; Pred. No. 1.5e-136;
Matches 684; Conservative 0; Mismatches 7; Indels 4; Gaps 4;
QY 47 gactgaagtagcaataatcatctcttcaaatatcatccgacacacacatgattca 106
DB 1 gactgaagtagcaataatcatctcttcaaatatcatccgacacacacatgattca 60
QY 107 agcaccgcacacagtcgcgggttttgggtatttcttcacgaataatcactgcttcagaatat 166
DB 61 agcaccgcacacagtcgcgggttttgggtatttcttcacgaataatcactgcttcagaatat 120
QY 167 gactgcacagaaatttcagcacagacattttcaactcaaaagcccttgcaaaaattattt 226
DB 121 gactgcacagaaatttcagcacagacattttcaactcaaaagcccttgcaaaaattattt 180
QY 227 gctagaaaaatgaaatcttagggactatccagatccctgtttggaattatgaccttttct 286
DB 181 gctagaaaaatgaaatcttagggactatccagatccctgtttggaattatgaccttttct 240
QY 287 ttgtagttatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 346
DB 241 ttgtagttatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 300
QY 347 ctttcagatattcattctgggctctgtttgtttgtttgtttgtttgtttgtttgtttgttt 406
DB 301 ctttcagatattcattctgggctctgtttgtttgtttgtttgtttgtttgtttgtttgttt 360
QY 407 gcagtgaagaaaaaacacagaaactctgataattatggccgaataatgaattttctt 466
DB 361 gcagtgaagaaaaaacacagaaactctgataattatggccgaataatgaattttctt 420
QY 467 agtgcctggggagcaatagctgggaatcattctcttcacatttg-gtttcactcagatca 525
DB 421 agtgcctggggagcaatagctgggaatcattctcttcacatttgagttttcactcagatca 480
QY 526 aaactacattgtgttattctcacaataatagtcagtgtaagcgttactgtcctgtt 585

DB 481 aaactacattgtgttattctcacaataatagtcagtgtaagcgttactgtcctgtat 540
QY 586 cttgggaatttgattacattgatgacttccagcat-tattgaattattcat-ttctctg 643
DB 541 cttgggaatttgattacattgatgacttccagcat-tattgaattattcat-ttctctg 600
QY 644 ctttt-ctcaatttgggggtgccactcagagatgtgattgttgcaaatgttgttact 702
DB 601 cttttactcaatttgggggtgccactcagagatgtgattgttgcaaatgttgttact 660
QY 703 agcactgtgagaataaagatgtgttaaaatctcaa 737
DB 661 agcactgtgagaataaagatgtgttaaaatctcaa 695
RESULT 7
AAS04279
ID AAS04279 standard; cDNA; 603 BP.
XX
AC AAS04279;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #1 cDNA sequence.
XX
KW Human; membrane protein; membrane receptor; IgE receptor; CD20;
KW physiological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..603
FT /tag= a
FT /product= "Human membrane protein #1"
FT /transl_except= (pos:595..600,aa:Cys)
XX
PN WO200146417-A2.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA;
XX
DR WPI; 2001-408646/43.
DR P-PSDB; AAU01210.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
XX
PS Claim 1; Page 29; 32pp; English.
XX
CC The present sequence encodes for novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IgE receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 603 BP; 162 A; 125 C; 104 G; 212 T; 0 other;

DE Novel human membrane protein #4 cDNA sequence.
 XX Human; membrane protein; membrane receptor; IgE receptor; CD20;
 KW physiological disorder; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 1..450
 CDS /tag- a
 FT /product= "Human membrane protein #4"
 XX WO200146417-A2.
 FN 28-JUN-2001.
 XX 12-DEC-2000; 2000WO-US33742.
 XX 22-DEC-1999; 99US-0171567.
 PR (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Turner CA;
 XX WPI; 2001-408646/43.
 DR P-PSDB; AAU01213.
 XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -
 XX Disclosure; Page 31; 32pp; English.
 XX The present sequence encodes for novel human membrane protein #4.
 CC Human membrane protein #4 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IgE receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 XX Sequence 450 BP; 120 A; 94 C; 77 G; 159 T; 0 other;
 SQ

Query Match 45.7%; Score 347.4; DB 22; Length 450;
 Best Local Similarity 97.0%; Pred. No. 1.9e-70;
 Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 98 atggattcaagcgcacagtcgcggtgttctgttattctccagaaatcactgct 157
 DB 1 atggattcaagcgcacagtcgcggtgttctgttattctccagaaatcactgct 60
 QY 158 tcagatatgagtcacagaactttcagcacgaccttttcaactcaagcccttgcaa 217
 DB 61 tcagatatgagtcacagaactttcagcacgaccttttcaactcaagcccttgcaa 120
 QY 218 aaattattgtcagaaaaatgaaatcttagggactatccagatccctgttgggaattatg 277
 DB 121 aaattattgtcagaaaaatgaaatcttagggactatccagatccctgttgggaattatg 180
 QY 278 accctttctttggaggtatcttctcttttccattgtgttaaacatcatccagggttccc 337
 DB 181 accctttctttggaggtatcttctcttttccattgtgttaaacatcatccagggttccc 240
 QY 338 ttatatattcttcaggatccattctgggctctgtttgtttcattatcaggagcc 397
 DB 338 ttatatattcttcaggatccattctgggctctgtttgtttcattatcaggagcc 397

DB 241 ttatatattcttcaggatccattctgggctctgtttgttcttaattctggagcc 300
 QY 398 ttcttaattgcagtgaaaagaaaaccacagaaactctgataatttgagccggaataatg 457
 DB 301 ttcttaattgcagtgaaaagaaaaccacagaaactctgggaattttgattacattgatg 360
 QY 458 aattt 462
 DB 361 acttt 365
 RESULT 10
 AAS04281
 ID AAS04281 standard; cDNA; 417 BP.
 XX
 AC AAS04281;
 XX 26-SEP-2001 (first entry)
 DT Novel human membrane protein #3 cDNA sequence.
 DE Human; membrane protein; membrane receptor; IgE receptor; CD20;
 KW physiological disorder; ss.
 KW Homo sapiens.
 OS Key Location/Qualifiers
 FH 1..417
 CDS /tag- a
 FT /product= "Human membrane protein #3"
 XX WO200146417-A2.
 XX 28-JUN-2001.
 XX 12-DEC-2000; 2000WO-US33742.
 XX 22-DEC-1999; 99US-0171567.
 PR (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Turner CA;
 XX WPI; 2001-408646/43.
 DR P-PSDB; AAU01212.
 XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -
 XX Disclosure; Page 30; 32pp; English.
 XX The present sequence encodes for novel human membrane protein #3.
 CC Human membrane protein #3 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IgE receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 XX Sequence 417 BP; 115 A; 91 C; 69 G; 142 T; 0 other;
 SQ

Query Match 44.8%; Score 340.2; DB 22; Length 417;
 Best Local Similarity 99.1%; Pred. No. 8.3e-69;
 Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 98 atggattcaagcaccgcacagtcggtgtttctgtgtattctccagaaatcactgct 157
|||||
Db 1 atggattcaagcaccgcacagtcggtgtttctgtgtattctccagaaatcactgct 60
|||||
QY 158 taagaatagtcacagaacttcagccagccacttttcaactcaaaagcccttgcaa 217
|||||
Db 61 taagaatagtcacagaacttcagccagccacttttcaactcaaaagcccttgcaa 120
|||||
QY 218 aaattattgctagaaaaatgaaatcttagggactatccagatcctgtttggaattatg 277
|||||
Db 121 aaattattgctagaaaaatgaaatcttagggactatccagatcctgtttggaattatg 180
|||||
QY 278 accttttttggaggtattctctcttcttcaactttgtttaaaccatataccaaagtttccc 337
|||||
Db 181 accttttttggaggtattctctcttcttcaactttgtttaaaccatataccaaagtttccc 240
|||||
QY 338 ttatattcttcagagatatacttctgggctctgttttgcataattcttgagacc 397
|||||
Db 241 ttatattcttcagagatatacttctgggctctgttttgcataattcttgagacc 300
|||||
QY 398 ttctaatgtagtgaagaaagaaacccacagaaactctgataata 442
|||||
Db 301 ttctaatgtagtgaagaaagaaacccacagaaactctgataata 345
|||||

RESULT 11

AAS04280

ID AAS04280 standard; cDNA; 246 BP.

XX AC AAS04280;

XX DT 26-SEP-2001 (first entry)

XX DE Novel human membrane protein #2 cDNA sequence.

XX KW Human; membrane protein; membrane receptor; IgE receptor; CD20;

XX KW physiological disorder; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..246

XX FT /*tag= a

XX FT /product= "Human membrane protein #2"

XX PN WO200146417-A2.

XX PD 28-JUN-2001.

XX PF 12-DEC-2000; 2000WO-US33742.

XX PR 22-DEC-1999; 99US-0171567.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Turner CA;

XX XX WPI; 2001-408646/43.

XX DR P-PSDB; AAU01211.

XX PT Polynucleotide encoding novel human membrane protein, useful for

XX PT identifying agonist, antagonist or modifiers or for producing

XX PT antibodies useful in therapeutic, diagnostic and pharmacogenomic

XX PT applications -

XX XX Disclosure; Page 30; 32pp; English.

XX CC The present sequence encodes for novel human membrane protein #2.

XX CC Human membrane protein #2 is 1 of 4 human membrane proteins

XX CC (AAU01210-AAU01213) given in the present invention. These membrane

XX CC proteins share structural similarity with membrane receptors such as

XX CC the IgE receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX SQ Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;

Query Match 31.9%; Score 242.8; DB 22; Length 246;

Best Local Similarity 99.2%; Pred No. 1.6e-46;

Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 455 atgaattcttagtgcctggagcaatagctggaatcattctcctcacatttggttc 514

|||||

Db 1 atgaattcttagtgcctggagcaatagctggaatcattctcctcacatttggttc 60

|||||

QY 515 atcctagatcaaaactacatttggttattctcaccacaaatagtcagtgtaaggctgtt 574

|||||

Db 61 atcctagatcaaaactacatttggttattctcaccacaaatagtcagtgtaaggctgtt 120

|||||

QY 575 actgtcctgttctgggaattttgattacattgatgactttcagcattatgaattattc 634

|||||

Db 121 actgtcctgttctgggaattttgattacattgatgactttcagcattatgaattattc 180

|||||

QY 635 atttctcgtcctttctcaattttgggggcccactcagaggattgtgattgaaacaatgt 694

|||||

Db 181 atttctcgtcctttctcaattttgggggcccactcagaggattgtgattgaaacaatgt 240

|||||

QY 695 tgttga 700

|||||

Db 241 tgttga 246

RESULT 12

AAH64745

ID AAH64745 standard; cDNA; 468 BP.

XX AC AAH64745;

XX DT 11-SEP-2001 (first entry)

XX DE Human secreted protein cDNA, SEQ ID NO: 21.

XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

XX KW GENSET; ss.

XX OS Homo sapiens.

XX PN WO200142451-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-IB01938.

XX PR 08-DEC-1999; 99US-0169629.

XX PR 06-MAR-2000; 2000US-0187470.

XX XX (BEST) GENSET.

XX PI Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX DR WPI; 2001-367870/38.

XX DR P-PSDB; AAG89142.

XX PT Full length GENSET human nucleic acids encoding potentially secreted

XX PT proteins, useful in gene therapy and vaccination against a variety of

XX PT diseases, and for diagnosis of those diseases -

XX PS Claim 7; Page 581; 921pp; English.

XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patients own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.

XX Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;

Query Match 30.7%; Score 233.4; DB 22; Length 468;
Best Local Similarity 80.5%; Pred. No. 2.6e-44;
Matches 273; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 23 agcaacaagaaaagacatggtctagactgaagtacacaaactaaatctctcttcaaat 82
Db 1 agcaacaagaaaagacatggtctagactgaagtacacaaactaaatctctcttcaaat 60
Qy 83 atcaccgacaccatcattggttcaagcaccgacacagtcgcggtgttcttggtatttcc 142
Db 61 atcaccgacaccatcattggttcaagcaccgacacagtcgcggtgttcttggtatttcc 120
Qy 143 ccagaaatcaactgcttcagaaatgagtcacagaaacttcacgacgacctttcaact 202
Db 121 ccagaaatcaactgcttcagaaatgagtcacagaaacttcacgacgacctttcaact 180
Qy 203 caaagcccttgcaaaaattatttgcagaaaaatgaaatctttagggactccagatc 262
Db 181 caaagcccttgcaaaaattatttgcagaaaaatgaaatctttagggactccatct 240
Qy 263 ctgtttggaatatgacctttcttttggagttatctctcttcaacttggtaaaacca 322
Db 241 ggggctgtgttttgcatttaattcttgagccttcctcaattgcagtgaaagaaacca 300
Qy 323 tatccaaagtttcccttcttatatttcttcaggatatacca 361
Db 301 cagaaactctgggaatttctgattacattgatgactttca 339

RESULT 13
AAAL6693
ID AAAL6693 standard; cDNA; 1330 BP.
XX
AC AAAL6693;
XX
DT 16-JUN-2000 (first entry)
XX
DE Human secreted protein clone pe246_1 nucleotide sequence SEQ ID NO:151.
XX Human; secreted protein; immunestimulant; immunosuppressant; virucide;
KW antibacterial; antifungal; cytostatic; antinflammatory; dermatological;
KW antidiabetic; antiaesthetic; antiarthritic; antirheumatic; protozoacide;
KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;
KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
KW connective tissue disease; multiple sclerosis; erythematosis;
KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
KW insulin dependent diabetes mellitus; graft-versus-host-disease;
KW autoimmune inflammatory eye disease; allergy; ss.

OS Homo sapiens.
XX WO200009552-A1.
PN
XX 24-FEB-2000.
PD
XX 13-AUG-1999; 99WO-US18298.
PF
XX 14-AUG-1998; 98US-0096622.
PR 17-AUG-1998; 98US-0096815.
PR 04-SEP-1998; 98US-0099229.
PR 23-OCT-1998; 98US-0105368.
PR 08-JAN-1999; 99US-0115234.
PR 12-FEB-1999; 99US-0119931.
PR 18-FEB-1999; 99US-0120575.
PR 30-APR-1999; 99US-0132020.
PR 11-AUG-1999; 99US-0096622.
XX (GENY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI Wong GG, Clark HF, Fechtel K;
XX WPI; 2000-205979/18.
DR P-PSDB; AAY94973.
XX
XX New polynucleotides encoding secreted proteins, which may have e.g.
PT nutritional, chemokine, immune stimulating or suppressing,
PT hematopoiesis regulating, tissue growth, activin/inhibin
PT antiinflammatory or tumor inhibition activity
XX
XX Claim 160; Page 606-607; 641pp; English.
XX
CC AAAL6618 to AAAL6697 encode the human secreted proteins given in
CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,
CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,
CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans
CC and animals. The polynucleotides can be used as markers for tissues in
CC which the protein is preferentially expressed, as molecular weight
CC markers on Southern gels, and as chromosome markers or tags to identify
CC chromosomes or to map gene positions. The proteins can be used in the
CC treatment of immune deficiencies and disorders, such as severe combined
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
CC infections. These infections include human immunodeficiency virus (HIV),
CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
CC candidiasis. The proteins can be used to treat autoimmune disorders such
CC as connective tissue disease, multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
CC autoimmune inflammatory eye disease. The proteins can also be used to
CC treat allergic conditions, such as asthma. AAAL6698 to AAAL6774 represent
CC probes for the human secreted proteins from the present invention.
XX
XX Sequence 1330 BP; 364 A; 285 C; 309 G; 372 T; 0 other;

Query Match 11.4%; Score 86.4; DB 21; Length 1330;
Best Local Similarity 51.1%; Pred. No. 1.7e-10;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;
Qy 216 aaaaattatttctgtagaaaaatgaaatcttagggactatccagatctgtttggaatta 275
Db 485 agaagtcttgaggaggagaaacccaaagctcttggttgaggatcttgactgcgcctga 544
Qy 276 tgacctttcttttggaggttatcttcttcttcaacttggtttaaacacatatccaggttgc 335
Db 545 tgagccttagcatggggaataacaatgatgtgtatggcatctaatctatgtgaagtaacc 604

FTXXPNXXPDXXPFXXPRXXPAXXPIXXDRXXPTXXPSXXCC

US5977072-A.
02-NOV-1999.
15-DEC-1998; 98US-0213389.
21-AUG-1997; 97US-0916902.
(INCYTE-) INCYTE PHARM INC.
Bandman O, Corley NC, Lal P;
WPI; 2000-012123/01.
P-PSDB; AAY50174.
High affinity immunoglobulin E receptor-like protein useful for
stimulating allergic and immune responses -
Example 5; Fig 1; 29pp; English.

570ctgttaactgtctctgtttcttgggaattttgattacattgatgaacttcagcaattattgaat629
686ggactatgtccatcttaataatgggtctctggatggcatgggtgctctcttaagtgtgctggaat745
630tattcaattctctgcctcttctcaatttttgggggtg663
746tctgcattgctgtgtccctctctgcctttggatg779

Search completed: August 30, 2002, 02:46:09
Job time: 4347 sec

FTXXPNXXPDXXPFXXPRXXPAXXPIXXDRXXPTXXPSXXCC

High affinity immunoglobulin E receptor-like protein useful for
stimulating allergic and immune responses -
Example 5; Fig 1; 29pp; English.

This sequence represents cDNA encoding human high affinity immunoglobulin
E (IgE) receptor-like protein (IGERB). Nucleic acids encoding IGERB were
initially identified in a brain cDNA library, this sequence being a
consensus. An allergic response is initiated by release of IgE
from B lymphocytes. The IgE molecules then bind to the high affinity
IgE receptor (FCRI) present on mast cells and basophils, which triggers
the release of histamine and proteases from cytoplasmic granules and
leads to the synthesis of effectors of the allergic and inflammatory
response, such as prostaglandins, leukotrienes and cytokines. As IGERB
binds IgE, it may be administered to stimulate allergic and immune
responses in patients in whom IGERB is under expressed or inactive and
to supplement the patients own production of the protein. IGERB may also
be used as an antigen for the production of antibodies and to identify
candidate agonists and antagonists of IGERB expression and activity.
The antibodies may also be used in diagnosis. Antibodies and antagonists
may be administered to downregulate IGERB activity and reduce the
potency of inflammatory and allergic responses. They may be used in this
way to treat inflammatory disorders such as multiple sclerosis,
osteoarthritis, asthma and some complications of cancer. Conversely, the
agonists may be used to enhance immune responses.

Sequence 1670 BP; 481 A; 350 C; 334 G; 501 T; 4 other;

Query Match 11.4%; Score 86.4; DB 21; Length 1670;
Best Local Similarity 51.1%; Pred. NO. 1.8e-10;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

QY 216 aaaaattattgtcagaataatgaaatcttaggactatccagatccctgtttggaatta 275
DB 326 agaagtcttgaggagagaccacaaagtccttgggtgtgcagattctgactgccctga 385
QY 276 tgacctttcttttgagttattctcttcttcaattttaaaccatatacgaagtttc 335
DB 386 tgagccttagcatgggaataaacaatgatgtatgcatcctaatactattggaagtaacc 445
QY 336 cctttatattcttcaggaatccattctgggctctgtttgttcatttaattctggag 395
DB 446 ctatttcocgtgtatcgggtgacacaaatttgggggtccagtaattgtttatttcaggat 505
QY 396 ccttcctaattgcagtgaagaagaaacacacagaactctgataatattgagccgaataa 455
DB 506 ccttgcaattgcagcagaattagaactacaaaagcctgtcccgaggtagtctaggaa 565
QY 456 tgaattttcttagtgcctgggggaataagctgggaatcattctcctcacatttggttca 515
DB 566 tgaatatcaccagctctgtactgctgcgcatcagggatcttaatacacacattagcttg 625
QY 516 tcttagat-----caaaactacatttgggttattctcaccacaaatagtcagtgtaagg 569
DB 626 cgttttattccatccacccttactgttaactactatggcaactcaataattgtcatg 685

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2002, 00:42:07 ; Search time 1960.86 Seconds
(without alignments)
8110.813 Million cell updates/sec

Title: US-09-821-821-1
Perfect score: 760
Sequence: 1 ttccagtgtccaggcagcc.....aaaaaaaaaaaaaaaaaaaaa 760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues 3595312
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htgo_inv.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| ----- | | | | | |

| | | | | | |
|----|-------|-------|--------|----|-----------|
| 1 | 760 | 100.0 | 760 | 6 | AX268522 |
| 2 | 689.2 | 90.7 | 694 | 9 | AF237907 |
| 3 | 687.4 | 90.4 | 697 | 9 | AF321127 |
| 4 | 686.2 | 90.3 | 691 | 6 | AB013103 |
| 5 | 666.4 | 87.7 | 676 | 6 | AX179789 |
| 6 | 598.2 | 78.7 | 603 | 6 | AX179781 |
| 7 | 377.2 | 49.6 | 382 | 6 | AX335001 |
| 8 | 347.4 | 45.7 | 450 | 6 | AX179787 |
| 9 | 340.2 | 44.8 | 417 | 6 | AX179785 |
| 10 | 250.4 | 32.9 | 138097 | 2 | AC027787 |
| 11 | 250.4 | 32.9 | 138097 | 2 | AC027787 |
| 12 | 250.4 | 32.9 | 147788 | 2 | AC015840 |
| 13 | 250.4 | 32.9 | 161238 | 9 | AP003127 |
| 14 | 250.4 | 32.9 | 166793 | 2 | AC090401 |
| 15 | 250.4 | 32.9 | 167934 | 9 | AP001034 |
| 16 | 242.8 | 31.9 | 246 | 6 | AX179783 |
| 17 | 88 | 11.6 | 1076 | 9 | AF068288 |
| 18 | 86.4 | 11.4 | 720 | 9 | AB022821 |
| 19 | 86.4 | 11.4 | 916 | 9 | AB013102 |
| 20 | 86.4 | 11.4 | 1550 | 9 | BC020648 |
| 21 | 86.4 | 11.4 | 1619 | 9 | AF237912 |
| 22 | 86.4 | 11.4 | 1669 | 6 | AR035695 |
| 23 | 86.4 | 11.4 | 1669 | 6 | AR083580 |
| 24 | 83.2 | 10.9 | 959 | 9 | AF350500 |
| 25 | 80.2 | 10.6 | 2805 | 9 | AF367473 |
| 26 | 64 | 8.4 | 1222 | 4 | SSR236932 |
| 27 | 63.6 | 8.4 | 2997 | 9 | AK057418 |
| 28 | 62.6 | 8.2 | 522 | 9 | AF354928 |
| 29 | 61.8 | 8.1 | 708 | 6 | AX101306 |
| 30 | 61.8 | 8.1 | 1010 | 10 | MUSFCERB |
| 31 | 61.6 | 8.1 | 1646 | 9 | HUMIERB |
| 32 | 61.6 | 8.1 | 1661 | 6 | AG8627 |
| 33 | 61.6 | 8.1 | 1661 | 6 | AR082014 |
| 34 | 61.6 | 8.1 | 1661 | 6 | BD003343 |
| 35 | 61.6 | 8.1 | 1661 | 6 | I25809 |
| 36 | 60 | 7.9 | 1628 | 9 | BC008487 |
| 37 | 58.2 | 7.7 | 1474 | 6 | I07272 |
| 38 | 58.2 | 7.7 | 1474 | 6 | I07277 |
| 39 | 58.2 | 7.7 | 1476 | 6 | HSCD20 |
| 40 | 58.2 | 7.7 | 1482 | 9 | BC002807 |
| 41 | 58.2 | 7.7 | 1597 | 6 | AX333252 |
| 42 | 58.2 | 7.7 | 1597 | 9 | HSCD20A |
| 43 | 57.2 | 7.5 | 899 | 10 | AF237910 |
| 44 | 57 | 7.5 | 1163 | 10 | AF237915 |
| 45 | 55 | 7.2 | 1146 | 9 | HUMBILYM |

ALIGNMENTS

| | | | | | | |
|------------|---|-----------------------------------|--------|-----|--------|-----------------|
| RESULT 1 | AX268522 | Sequence 1 from Patent WO0174903. | 760 bp | DNA | linear | PAT 29-OCT-2001 |
| LOCUS | AX268522 | Sequence 1 from Patent WO0174903. | | | | |
| DEFINITION | AX268522 | Sequence 1 from Patent WO0174903. | | | | |
| ACCESSION | AX268522.1 | GI:16541676 | | | | |
| VERSION | human. | | | | | |
| KEYWORDS | human. | | | | | |
| SOURCE | human. | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | |
| REFERENCE | 1 (sites) | | | | | |
| AUTHORS | Welcher,A.A. and Calzone,F.J. | | | | | |
| TITLE | Cd20/ige-receptor like molecules and uses thereof | | | | | |
| JOURNAL | Patent: WO 0174903-A 1 11-OCT-2001; | | | | | |
| FEATURES | Amgen, Inc. (US) | | | | | |
| source | Location/Qualifiers | | | | | |
| | 1..760 | | | | | |
| | /organism="Homo sapiens" | | | | | |
| | /db_xref="taxon:9606" | | | | | |
| | 98..700 | | | | | |
| CDS | /note="unnamed protein product" | | | | | |
| | /codon_start=1 | | | | | |

/protein_id="CAD10310.1"
/db_xref="GI:16541677"
/translation="MDSSTAHSPVFLVFPPEITASEVESPELSATFSTQSPLOKLFA
RKMKTLGTLQILGIMTFPSGVIFLTLKPYPRPFIFLSGVPWGSVLFNSGAF
LAVKPKTETLILSRINWFLSALGATAGIILLTGFILDONVICYSHNSOCKRAVT
VLFGLITLMTFSLIILFSLPSILGCHSEDCCEQC"
232 a 157 c 126 g 245 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 760; DB 6; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.2e-161;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ttccagtgtccaggcgtccagcagcaagaaagaaacatggtctagactgaagtaccaa 60
Db 1 TTCCAGTGTCTCCAGGAGCCTCAGCACAGAAAGAAACATGCTAGACTGAAGTACCAA 60
Qy 61 ctaaatcatctcttcaaatatcaccagcaccatcatggtattcaagcaccgacacag 120
Db 61 CTAAATCATCTCCTTTCAAAATATCACGACACCATCATGGATTCAGACCCGACACAG 120
Qy 121 tccggtgttcttggtatttccctccagaatacactgcttcacagaatatgagtcacagaact 180
Db 121 TCCGGTGTCTTCTGGTATTTCTCCAGAAATCACTGCTTCAGAAATATGAGTCCACAGAACT 180
Qy 181 ttccagcagacactttcaactcaagcccttcgaaataatttctagaaataatgaa 240
Db 181 TTCCAGCACGACCTTTCAACATCAAGCCCTTCGAAATATTTCTAGAAATATGAA 240
Qy 241 aattcagggaactaccagactgctgttggaattatgacctttcttttggagttactct 300
Db 241 AATCTTAGGACTATCCAGATCTGTTTGGAAATATGAGCTTTCTTTTGGAGTTACTT 300
Qy 301 cctttcactttgttaaaaccataccaggttcccttcttatattttttcaggatatcc 360
Db 301 CCTTTTCACTTTGTHAAACCAATACCAAGGTTTCCCTTTATATTTCTTCAGAGATACC 360
Qy 361 attctggggctctgtttgttcattaatctggagccttcctaatgtagtgaagaa 420
Db 361 ATTCTGGGGCTCTGTGTTTTCATTAATCTGGAGCCTTCCTAATTCAGTGAAGAA 420
Qy 421 aaccacagaactctgataataattgagccgaataatgatttcttagtccctggagc 480
Db 421 AACACAGAAACTCTGATTAATATGAGCGGAATATGAAATTTCTTAGTCCCTGGGAGC 480
Qy 481 aatagctggaatcaattctctccacatttgggtttcattcctagatcaaaactacattgtg 540
Db 481 AATAGCTGGAATCATTTCTCCTCACATTTTGGTTTCATCTCTAGATCAAAACTACATTTGTG 540
Qy 541 ttattctccaaaatagtacagtgaagcgttactgtctctgtctctgttggaatttggat 600
Db 541 TTATTCACCAAAATAGTCAGTGAAGGCTGTACTGCTCTGTTCTTGGAAATTTGAT 600
Qy 601 tacattgagactttccagcattatgaattatcaattctctgccttctcaattttggg 660
Db 601 TACATTTGATGACTTTTCCAGCATTTATGAATATTCATTTCTCTGCTTTCTCAATTTTGG 660
Qy 661 gtccactcagagattgattggaacaaatggtttgtagtagcactgtgagaataaag 720
Db 661 GTGCCACTCAGAGATTGTGATTGTGAACAAATGTGTTGTGACTAGCCTGTGAGAAATAAG 720
Qy 721 atgtgttaaaatctcaaaaaaataaaaaaataaaaaa 760
Db 721 ATGTGTTAAATCTCAAAAAAATAAAAAAATAAAAAA 760

RESULT 2

AF237907 AF237907 694 bp mRNA linear PRI 17-APR-2001
LOCUS Homo sapiens MS4A5 protein mRNA, complete cds.
DEFINITION AF237907
ACCESSION AF237907
VERSION AF237907.1 GI:13649400

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

Liang, Y. and Tedder, T.F.
1 (bases 1 to 694)
Identification of a cd20-, fcepsilonribeta-, and htm4-related gene
family: sixteen new ms4a family members expressed in human and
mouse

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

Genomics 72 (2), 119-127 (2001)

Submitted (24-FEB-2000)

Direct Submission

Dr., Durham, NC 27710, USA

Location/Qualifiers

source

1. 694

/organism="Homo sapiens"

/db_xref="taxon:9606"

55..657

/codon_start=1

/product="MS4A5 protein"

/protein_id="AAK37416.1"

/db_xref="GI:13649401"

/translation="MDSSTAHSPVFLVFPPEITASEVESPELSATFSTQSPLOKLFA
RKMKTLGTLQILGIMTFPSGVIFLTLKPYPRPFIFLSGVPWGSVLFNSGAF
LAVKPKTETLILSRINWFLSALGATAGIILLTGFILDONVICYSHNSOCKRAVT
VLFGLITLMTFSLIILFSLPSILGCHSEDCCEQC"

BASE COUNT 198 a 143 c 116 g 237 t

ORIGIN

Query Match

Best Local Similarity 90.7%; Score 689.2; DB 9; Length 694;

Matches 691; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 44 ctgagctgaagtaccacataatcatctctcttcaaaattatcacccagacacatcatggt 103

Db 1 CTAGACTGAAGTACCACATAAATCATCTCTCTTCAAAATATCACCCGACACCATCATGGAT 60

Qy 104 tcaagcaccgacacagtcggtgtttctggtattctctccagaaatcactgcttcagaa 163

Db 61 TCAGACACCCACACACTCGGTGTTCTGGTATTTCTCCAGAAATCAGTCGCTTCAGAA 120

Qy 164 tatgagtcacagaaactttcagccagacacttttcaactcaaaagcccttgcaaaaaa 223

Db 121 TATGAGTCCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCAAAAATTA 180

Qy 224 ttgtctagaaaaatgaaatcttagggactatccagatcctgtttggaattatgaccttt 283

Db 181 TTGTCTAGAAAAATGAAATCTTAGGGACTATCCAGATCTCTTTGGAAATATGACCTTT 240

Qy 284 tctttggagttatctctcttcttcaactttgttaaaacacatccaaaggtttccctttata 343

Db 241 TCCTTTGGAGTTATCTCTCTTTTCACTTTGTTAAAAACCATCCAAAGGTTTCCCTTTATA 300

Qy 344 ttctttcagagatatccattctggggctgtgtttgttcatttaattcttgagccttcta 403

Db 301 TTTCTTTTCAGGATATCCATCTCGGGGCTCTGTTTGTTCATTAATTTCTGGAGCCTTCTCTA 360

Qy 404 attgagtgaaaaagaaaaaacacagaaactctgataaattatgagccgaaataaataattt 463

Db 361 ATTGCAGTGAAGAAAAAACACACAGAAACTCTGTATATATTTGAGCCGAATAATGAATTT 420

Qy 464 cttagtgcctggggagcaatagctgggaatcattctccctcacatttggtttcatcctagat 523

Db 421 CTTAGTGCCTGGGAGCAATAGCTGGAATCATTTCTCCATCATTTGGTTTCATCCTTAGAT 480

Qy 524 caaaactacatttgggttattctcaccacaaatagtcaggtgtaaggtgttactgtcctg 583

Db 481 CAAAACTACATTTGTGTTATTCTCACCACCAAAATAGTCAGTGTGAAGGCTGTTACTGTCTG 540

| | | | |
|------------|---|---|--------|
| Qy | 162 | aatatgagtcacagaactttcagccagcaacttttcaactcaaaagcccttgcaaaaat | 221 |
| Db | 121 | AATATGAGTCACAGAAACTTTACGCCACGACCTTTTCAACTCAAAAGCCCTTGCAGAAAT | 180 |
| Qy | 222 | tatttgcataaaaaataaaacttagggactaccagactcctgttgggaattatgaact | 281 |
| Db | 181 | TATTTGCTAGAAAATGAAATCTTAGGGACTATCCAGATCCTGTTTGGAAATATGACCT | 240 |
| Qy | 282 | ttcttttggagttatcttctcttctacttggttataaaccatatccaaggtttcccttta | 341 |
| Db | 241 | TTCTTTTGGAGTTATCTTCCTTTTCACTTGTGTAAACCATATCCAAGGTTTCCCTTTA | 300 |
| Qy | 342 | tatttcttcaggatatccattcttggggctctgttttgttcataaattctggagccttcc | 401 |
| Db | 301 | TATTTCTTTTCAGGATATCCATTTCTGGGGCTCTGTTTGTGTCATTAAATCTCGAGCCCTTCC | 360 |
| Qy | 402 | taattgcagtgaagaagaaacacacagaaactctgataaattagagccgaataatgaatt | 461 |
| Db | 361 | TAATTCAGTGAANAAGAAACACACAGAAACTCTGATAATATTGACCGAATATGAATT | 420 |
| Qy | 462 | ttcttagtgccttggggagcaatagctggaatcatctctctcacatttgggttctcctag | 521 |
| Db | 421 | TTCTTAGTGCCCTGGGAGCAATAGCTGGAATCATTTCTCTCACATTTGGTTTCATCCTAG | 480 |
| Qy | 522 | atcaaaactacatttgggttattctccacaaaatagtcagtgtaagcgttactgtcc | 581 |
| Db | 481 | ATCAAAACTACATTTGTTGGTTATTTCTCACCAAAATATGTCAGTGAAGGCTGTTACTGTCC | 540 |
| Qy | 582 | tgtcttgggaatttggattacattgatgacttccagcattatgaattcattcttc | 641 |
| Db | 541 | TGTTGTTGGGAATTTGATACATGATGACTTTCAGCATTTATTGAATTAATCATTTCTC | 600 |
| Qy | 642 | tgctttctcaatttgggggtgccactcagaggattgtgtgtgtgaaacaattgtgtgac | 701 |
| Db | 601 | TGCTTTCTCAATTTGGGGTGCCACTCAGAGGATTTGATTGTGAACAATGTTGTTGCAC | 660 |
| Qy | 702 | tagcactgtgagaataaagatgtgttaaaatctcaaa | 738 |
| Db | 661 | TAGCAGTGTGAGATAAGATGTGTAAATAAATAA | 697 |
| RESULT 4 | | | |
| AB013103 | 4 | 591 bp | linear |
| LOCUS | AB013103 | 591 bp | linear |
| DEFINITION | Homo sapiens mRNA for MS4A5, complete cds. | | |
| ACCESSION | AB013103 | | |
| VERSION | AB013103.1 | GI:11559213 | |
| KEYWORDS | MS4A5; CD20-like 2. | | |
| SOURCE | Homo sapiens testis cDNA to mRNA. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | Ishibashi,K., Suzuki,M., Sasaki,S. and Imai,M. | | |
| TITLE | Identification of a new multigene four-transmembrane family (MS4A) related to CD20, HTm4 and beta subunit of the high-affinity IgE receptor | | |
| JOURNAL | Gene 264 (1), 87-93 (2001) | | |
| MEDLINE | 21142397 | | |
| REFERENCE | 2 (sites) | | |
| AUTHORS | Ishibashi,K., Sasaki,S. and Marumo,F. | | |
| TITLE | Cloning of three CD20 homolog from human, putative calcium channels. | | |
| JOURNAL | Unpublished | | |
| REFERENCE | 3 (bases 1 to 691) | | |
| AUTHORS | Ishibashi,K. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo 113-8519, Japan (E-mail:kishibashi.med2@med.tmd.ac.jp, Tel:81-3-5803-5223, Fax:81-3-5803-0132) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..691 | | |

/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translation="MSSTAHSPVFLVPPPEITASVESTELSATTFSTQSPLOKFLA
RKMKILGIIQLRIGIMTFSEGVIFLTLKPYRPFIFLSPYFPGSVLFINSGAFL
IAYRKTEFLIRLIMNPLSALGAIGAILTLFGILDLQNYICGYSQNSQCAVAT
VLFLGILITLMTSIIELFSLFSLFSLGHSIEDCDCEQC"
BASE COUNT 195 a 142 c 117 g 237 t
ORIGIN

Query Match 90.3%; Score 686.2; DB 9; Length 691;
Best Local Similarity 99.6%; Pred. No. 5.4e-145;
Matches 688; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 47 gactgaagtaccactaaatcatctcttccaaattatcccgacacacatcatggattca 106
Db 1 GACTGAAGTACCAACTAAGTCATCTCCTTTCAAATATATCACCGACACCATCGGATCA 60
Qy 107 agcaccgacacagctccgggtgttctggtatttctccagagaatcactgcttcagaatat 166
Db 61 AGCACCACACAGCTCCGGTGTCTGGTATTTCTCCAGAAATCAGTGTTCAGAAATAT 120
Qy 167 ggtccacagaaactttcagccacagacacttttcaactcaaaagcccttgcaaaaattatt 226
Db 121 GAGTCCACAGAACTTTTCAGCCACAGACCTTTTCACTCAAAAGCCCTTGCAAAAATATT 180
Qy 227 gctgaataatgaaacttttaggactatccagatcctgtttggaattatgacatttct 286
Db 181 GCTAGAAATGAAATCTTAGGAGACTATCCAGATCCTGTTTGGAAATATGACCTTTCT 240
Qy 287 ttggaggtattcttcttcaactttgttaaaccatcatccaagggttcccttttatatt 346
Db 241 TTTGGAGTTATCTTCTTTTCACTTTCTTAAACCATATCCAAAGTTTTCCTTTATATT 300
Qy 347 ctccagagatccattctggggtctgtttgttcaattatctcagagccttccctaatt 406
Db 301 CTTTCAGGATATCATTTCTGGGCTCTGTTTGTTCATTAATTTCTGGAGCCTTCTTAAT 360
Qy 407 gcagtgaagaagaaacacagaaactctgataattatggagccgaataatgaaatttctt 466
Db 361 CAGTGAAGAAAGAAACACAGAAACTCTGATAATATTGAGCCGAAATATGAATTTCTTT 420
Qy 467 agtgccctgggagcaatagctgggaatcattctcctcacatttgggttccatcctagatcaa 526
Db 421 AGTCCCTGGGAGCAATAGCTGGAATCATTTCTCCTACATTTGGTTTCATCCTTAGATCAA 480
Qy 527 aactacatttgggttattctcaccaaaatagtcagtgtaagcgttactgtcctgttc 586
Db 481 AACTACATTTGGGTATTCTCACCAAAATAGTCAGTGAAGCCTGTACTGTCCCTGTTTC 540
Qy 587 ttgggaatttggattacattgacttttcagcattattgaattattcatttctctgcct 646
Db 541 TTGGGAATTTGATTACATTTGATGACTTTTCAGCATTTATTGAATATTCAATTTCTCGCCT 600
Qy 647 ttctcaatttgggggtgccactcagaggattgtgatttgaaacaattgttggactagca 706
Db 601 TTCFCAATTTTGGGGTGCCTACTCAGAGGATTGTGATTGTGAACAATGTTGTTGACTAGCA 660
Qy 707 ctgtgagaataaagatgtgttaaaattctcaa 737
Db 661 CTGTGAGATAAAGATGTGTTAAATATAAA 691

RESULT 5

AX179789
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN

AX179789
Sequence 9 from Patent WO0146417.
AX179789
AX179789.1 GI:15132148
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Walke,D.W. and Turner,C.A.
Human membrane proteins and polynucleotides encoding the same
having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 9 28-JUN-2001;
Lexicon Genetics Incorporated (US)
Location/Qualifiers
1..676
/organism="Homo sapiens"
/db_xref="taxon:9606"
193 a 140 c 112 g 231 t

Query Match 87.7%; Score 666.4; DB 6; Length 676;
Best Local Similarity 99.1%; Pred. No. 1.6e-140;
Matches 670; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 66 tcattctcttcaaatattcacgcacacatcatggattccaagcaccgcacacagctccgg 125
Db 1 TCATCTCCTTTCAAATATATCACCGACACCATCATGGATTCAAGCACCCGACACAGTCGG 60
Qy 126 tgtttctggtatttctccagaaatcactgcttcagaatatgagtcacacagaaactttcag 185
Db 61 TGTTCCTGATTTCTTCAGAAATCAGTGTTCAGATATATGATGCTCCAGAACTTTTCAG 120
Qy 186 ccacgaccttttcaactcaaaagcccttgcaaaaattatttctcagagaaatgaaaatct 245
Db 121 CCACGACCTTTTCAACTCAAAAGCCCTTGCAAAAATATTATTTGCTAGAAAATGAAAATCT 180
Qy 246 taggactatccagatcctgttggaaatagacatttcttttggagattatcttcccttt 305
Db 181 TAGGACTATCCAGATCCTGTTTGGAAATATGACCTTTCTTTTGGAGTTATCTTCTCTT 240
Qy 306 tcactttgttaaaacatattccaaggtttccctttatatttcttccagatattccattct 365
Db 241 TCACCTTGTTAAACCATATCCAAAGTTTCCCTTTATATTCTTTTCAGGATATCCATTCT 300
Qy 366 ggggtctgtttgttcttaattatctcagagccttccctaatcagtgagtgaaagaaacca 425
Db 301 GGGGCTCTGTTTGTTCATTAATTTCTGGAGCCTTCTTAATTCAGTGAAGAAAGAAAACCA 360
Qy 426 cagaactcgaataattagccgaataatgaaatttctttagtgccctgggagcaatag 485
Db 361 CAGAAACTCTGATAATATTGAGCCGAATTAATGAATCTTCTTAGTGCCCTTGAGAGCAATAG 420
Qy 486 ctggaatcattctcctcacatttgggtttccctcctagatcctcaaaactacatttgggttatt 545
Db 421 CTGGAATCATTTCTCCTACATTTGGTTTTCATCCCTAGATCAAACTACATTTTGGTTATT 480
Qy 546 ctcaacaaaatagtcagtgtaagcgtgttaactgtcctgttcttgggaaatttggattacat 605
Db 481 CTCACAAAATAGTCAGTGAAGGCTGTACTGTCTCTCTTCTTGGAAATTTTGGATTACAT 540
Qy 606 tgatgacttcagcattattgaattattcatttctcctccttctcatttcttgggggtcc 665
Db 541 TGATGACTTTTCAGCAATATTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Qy 666 actcagaggattgtgattgtaacaaatgttctgactagcactgtgagaaataaagagtgtg 725
Db 601 ACTCAGAGGATGTGATTGTGAACAATTTGTTGACTAGCAGTGTGAGAAATTAAGATGTG 660
Qy 726 ttaaaatctcaaaaaa 741

```
||||| |||||
Db 661 TTAATAAATAAAAAA 676

RESULT 6
AX179781
LOCUS AX179781 603 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0146417.
ACCESSION AX179781
VERSION AX179781.1 GI:15132144
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 603)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
having homology to cd20 proteins and ige receptors
JOURNAL Patent: WO 0146417-A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..603
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 162 a 125 c 104 g 212 t
ORIGIN

Query Match 78.7%; Score 598.2; DB 6; Length 603;
Best Local Similarity 99.5%; Pred. No. 4e-125;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 98 atgattcaagcaccgacacagcccggtgtttctgtgatttctccagaaatcaactgct 157
Db 1 ATGATTCAAGCACCACAGTCCGGTGTTCGTGATTCTCTCCAGAAATCAGTCT 60

Qy 158 tcagaatatgagtcacagaaacttcagccagcagcactttcacaactcaaaccccttgc 217
Db 61 TCAGAAATATGAGTCACAGAACTTTCAGCCACGACCTTTTCAACTCAAGCCCTTGC 120

Qy 218 aaattatttctagaaataaataatttaggagactccagatccctgtttgaaattatg 277
Db 121 AAATTATTGCTAGAAATAAATAATTCTAGGAGACTATCCAGATCCCTGTGTGAAATATG 180

Qy 278 acctttcttttgagttatctctcttctcaattgtttaaacaatccaaaggtttccc 337
Db 181 ACCTTTCTTTGGAGTATCTCTCTTTTCACCTTGTATAACCATATCCAAAGGTTTCCC 240

Qy 338 ttatatttcttcaggatatccattctggtggtctctgttttgcatttaattctggagcc 397
Db 241 TTTATATTCTTTCAGGATATCCATTCTGGGGCTCTGTTTGTTCATTAATTCGGAGCC 300

Qy 398 ttcctaatttcagtgaaagaaacacacagaaactctgataatattgagccgaataatg 457
Db 301 TTCCTAATTCAGTGAAAGAAACACAGAAACTCTGATAATATTGAGCCGAAATATG 360

Qy 458 aattttcttagtcctgggagcaatagctggaatcattctcccaacttttggtttcactc 517
Db 361 AATCTCTCTTAGTCCCTCAGAGCAATAGCTGGAATCATCTCCCTCACATTTGGTTTCATC 420

Qy 518 ctagatcaaaactacatttggtttattctcaccaaaatagctagtgtaaggctgttact 577
Db 421 CTAGATCAAAACTACATTGTGGTATTCTCCACCAAAATAGTCAGTGAAGGCTGTACT 480

Qy 578 gtccctgttcttggaattttgattacattgatgacttcagcatttgatttattcatt 637
Db 481 GTCCCTGTCTCTGGGAATTTTGATTACATTGATGACTTTCACATATTGAATATTTCATT 540

Qy 638 tctctgcctttctcaatttttgggtggccactcagaggtattgtgattggaacaattgtg 697
Db 541 TCTCTGCCTTCTCAATTTTGGGGTGCCACTCAGAGGATTGTGATTGTGAACAATGTGT 600

Qy 698 tga 700
Db 601 TGA 603

RESULT 7
AX335001/c
LOCUS AX335001 382 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 5510 from Patent WO0194629.
ACCESSION AX335001
VERSION AX335001.1 GI:18125720
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 5510 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source 1..382
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 135 a 73 c 67 g 107 t
ORIGIN

Query Match 49.6%; Score 377.2; DB 6; Length 382;
Best Local Similarity 99.2%; Pred. No. 3.2e-75;
Matches 379; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 356 tatcaattctgggctctgtttgttcatttaattctggagccttctcaattgcagtgaaa 415
Db 382 TATCCATTCTGGGGCTCTGTTTGTTCATTAATCTGGAGCCCTCTTAATTGCAGTGAAA 323

Qy 416 agaaaaaccacagaaactctgataattgagccgaataataatgttctttagtgcctg 475
Db 322 AGAAAAACCACAGAAACTCTGATAATATTGAGCCGAATAATGAATTTCTTAGTGCCTG 263

Qy 476 ggaacaatagctggaatcattctcctcacatttggttctcactagatcaaaactacatt 535
Db 262 GGAGCAATAGCTGGAATCATTTCTCCATTTGGTTCATCTAGATCAAAACTACATT 203

Qy 536 tfggttattctcaccaaaatagctagtgtaaggctgttactgtctctgttgggaatt 595
Db 202 TGTGGTATTCTCACCAAAATAGTCAGTGTAGGCTGTACTGTCTGTCTTGTGGGAAT 143

Qy 596 ttgattacattgatgactttcagcattattgaaattattcattctctgcttctcatt 655
Db 142 TTGATTACATTGATGACTTTCAGCATTATTGAATATTTCATTCTCTGCCTTCTCAATT 83

Qy 656 ttgggtgacctcagagattgtgattgaaacaatgttgttgcactagcactgtgagaa 715
Db 82 TTGGGTGCCACTCAGAGGATTTGATTGTGAACAATTTGTGTGACTAGACTGTGAGAA 23

Qy 716 taaagatgtgtaaaatctcaa 737
Db 22 TAAAGATGTGTAAATATATAA 1

RESULT 8
AX179787
LOCUS AX179787 450 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 7 from Patent WO0146417.
ACCESSION AX179787
VERSION AX179787.1 GI:15132147
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 450)

AUTHORS

Walke,D.W. and Turner,C.A.

TITLE

Human membrane proteins and polynucleotides encoding the same

JOURNAL

having homology to cd20 proteins and ige receptors

PATENT

Patent: WO 0146417-A 7 28-JUN-2001;

LEXICON

Genetics Incorporated (US)

FEATURES

Location/Qualifiers

source

1. .450

BASE COUNT

120 a 94 c 77 g 159 t

ORIGIN

/organism="Homo sapiens"

/db_xref="taxon:9606"

Query Match

45.7% Score 347.4; DB 6; Length 450;

Best Local Similarity

97.0% Pred. No. 1.7e-68;

Matches

354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 98 atgattcaagcaccgacacacagtcggtttctgtgtatttctccagaaatcaactgct 157

|||||

Db 1 ATGGATTCAAGCACCACACAGTCGGTGTCTGTGATTTCTCCAGAAATCACTGCT 60

|||||

Qy 158 tcagaatatgattccacagaaacttccagccagcacttttcaactcaaaccccttgcaa 217

|||||

Db 61 TCAGAAATGAGTCCACAGAACTTTTCAGCCAGCAGCTTTTCAACTCAAAAGCCCTTGCAA 120

|||||

Qy 218 aaattattgctagaaaaatgaaactcttagggactatccagatccctgtttggaattatg 277

|||||

Db 121 AAATTATTGCTAGAAAATGAAATCTTAGGGACTATCCAGATCTGTTTGGAAATATG 180

|||||

Qy 278 acctttctttggagtattctcttcttcaactttgtttaaaccatataccaaggtttccc 337

|||||

Db 181 ACCTTTCTTTGGAGTTATCTTCTTCCCTTGTAAACCATATCAAGGTTTCCC 240

|||||

Qy 338 ttatattcttcagatatccattctgggtctgtttgttcaatttaattcttgagcc 397

|||||

Db 241 TTTATATTCTTTCAGGATATCCATCTCTGGGCTCTGTTTGTTCATTAATCTGGAGCC 300

|||||

Qy 398 ttctaatgtcagtgaaagaaacacacagaaactctgataatttgagccgaataatg 457

|||||

Db 301 TTCCTAATTCGAGTGAAGAAAGAAACCCAGAAACTCTGATCAAA 345

|||||

Qy 458 aattt 462

|||||

Db 361 ACTTT 365

RESULT 9

AX179785

LOCUS

Sequence 5 from Patent WO0146417. linear PAT 06-AUG-2001

DEFINITION

ACCESSION

AX179785

VERSION

AX179785.1

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Walke,D.W. and Turner,C.A.

TITLE

Human membrane proteins and polynucleotides encoding the same

JOURNAL

having homology to cd20 proteins and ige receptors

PATENT

Patent: WO 0146417-A 5 28-JUN-2001;

LEXICON

Genetics Incorporated (US)

FEATURES

Location/Qualifiers

source

1. .417

BASE COUNT

ORIGIN

115 a 91 c 69 g 142 t

Query Match 44.8% Score 340.2; DB 6; Length 417;

Best Local Similarity 99.1% Pred. No. 7.3e-67;

Matches 342; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 98 atgattcaagcaccgacacacagtcggtttctgtgtatttctccagaaatcaactgct 157

|||||

Db 1 ATGGATTCAAGCACCACACAGTCGGTGTCTTGGTATTTCTCCAGAAATCACTGCT 60

|||||

Qy 158 tcagaatatgattccacagaaacttccagccagcacttttcaactcaaaccccttgcaa 217

|||||

Db 61 TCAGAAATGAGTCCACAGAACTTTTCAGCCAGCAGCTTTTCAACTCAAAAGCCCTTGCAA 120

|||||

Qy 218 aaattattgctagaaaaatgaaactcttagggactatccagatccctgtttggaattatg 277

|||||

Db 121 AAATTATTGCTAGAAAATGAAATCTTAGGGACTATCCAGATCTGTTTGGAAATATG 180

|||||

Qy 278 acctttctttggagtattctcttcttcaactttgtttaaaccatataccaaggtttccc 337

|||||

Db 181 ACCTTTCTTTGGAGTTATCTTCTTCCCTTGTAAACCATATCAAGGTTTCCC 240

|||||

Qy 338 ttatattcttcagatatccattctgggtctgtttgttcaatttaattcttgagcc 397

|||||

Db 241 TTTATATTCTTTCAGGATATCCATCTCTGGGCTCTGTTTGTTCATTAATCTGGAGCC 300

|||||

Qy 398 ttctaatgtcagtgaaagaaacacacagaaactctgataata 442

|||||

Db 301 TTCCTAATTCGAGTGAAGAAAGAAACCCAGAAACTCTGATCAAA 345

|||||

RESULT 10

AC027787

LOCUS

DEFINITION

Homo sapiens

chromosome 15 clone RP11-196E16 map 15, LOW-PASS

SEQUENCE SAMPLING.

ACCESSION

AC027787

VERSION

AC027787.2

KEYWORDS

HTG; HTGS_PHASE0.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 15, clone RP11-196E16

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 138097)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Levine,R., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Klein,J., Liou,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., McElrath,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye.W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

Submitted (01-APR-2000)

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Aug 18, 2000 this sequence version replaced g1:7382629.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9008
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 664 763: gap of 100 bp
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length
* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 12682: gap of 100 bp
* 12683 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length

* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp

| | | | | | |
|---------|--------|----------------------------|---------|--------|----------------------------|
| * 16684 | 17376: | contig of 693 bp in length | * 45881 | 45980: | gap of 100 bp |
| * 17377 | 17476: | gap of 100 bp | * 45981 | 46683: | contig of 703 bp in length |
| * 17477 | 18178: | contig of 702 bp in length | * 46884 | 46783: | gap of 100 bp |
| * 18179 | 18278: | gap of 100 bp | * 46784 | 47481: | contig of 698 bp in length |
| * 18279 | 18962: | contig of 684 bp in length | * 47482 | 47581: | gap of 100 bp |
| * 18963 | 19062: | gap of 100 bp | * 47582 | 48290: | contig of 709 bp in length |
| * 19063 | 19673: | contig of 611 bp in length | * 48291 | 48390: | gap of 100 bp |
| * 19674 | 19773: | gap of 100 bp | * 48391 | 49077: | contig of 687 bp in length |
| * 19774 | 20436: | contig of 663 bp in length | * 49078 | 49177: | gap of 100 bp |
| * 20437 | 20536: | gap of 100 bp | * 49178 | 49879: | contig of 702 bp in length |
| * 20537 | 21233: | contig of 697 bp in length | * 49880 | 49979: | gap of 100 bp |
| * 21234 | 21333: | gap of 100 bp | * 49980 | 50684: | contig of 705 bp in length |
| * 21334 | 22042: | contig of 709 bp in length | * 50685 | 50784: | gap of 100 bp |
| * 22043 | 22142: | gap of 100 bp | * 50785 | 51507: | contig of 723 bp in length |
| * 22143 | 22808: | contig of 666 bp in length | * 51508 | 51607: | gap of 100 bp |
| * 22809 | 22908: | gap of 100 bp | * 51608 | 52320: | contig of 713 bp in length |
| * 22909 | 23595: | contig of 687 bp in length | * 52321 | 52420: | gap of 100 bp |
| * 23596 | 23695: | gap of 100 bp | * 52421 | 53133: | contig of 713 bp in length |
| * 23696 | 24391: | contig of 696 bp in length | * 53134 | 53233: | gap of 100 bp |
| * 24392 | 24491: | gap of 100 bp | * 53234 | 53933: | contig of 700 bp in length |
| * 24492 | 25190: | contig of 699 bp in length | * 53934 | 54033: | gap of 100 bp |
| * 25191 | 25290: | gap of 100 bp | * 54034 | 54703: | contig of 670 bp in length |
| * 25291 | 25981: | contig of 691 bp in length | | | |
| * 25982 | 26081: | gap of 100 bp | | | |
| * 26082 | 26765: | contig of 684 bp in length | | | |
| * 26766 | 26865: | gap of 100 bp | | | |
| * 26866 | 27563: | contig of 698 bp in length | | | |
| * 27564 | 27663: | gap of 100 bp | | | |
| * 27664 | 28377: | contig of 714 bp in length | | | |
| * 28378 | 28477: | gap of 100 bp | | | |
| * 28478 | 29184: | contig of 707 bp in length | | | |
| * 29185 | 29284: | gap of 100 bp | | | |
| * 29285 | 29983: | contig of 699 bp in length | | | |
| * 29984 | 30083: | gap of 100 bp | | | |
| * 30084 | 30775: | contig of 692 bp in length | | | |
| * 30776 | 30875: | gap of 100 bp | | | |
| * 30876 | 31565: | contig of 690 bp in length | | | |
| * 31566 | 31665: | gap of 100 bp | | | |
| * 31666 | 32360: | contig of 695 bp in length | | | |
| * 32361 | 32460: | gap of 100 bp | | | |
| * 32461 | 33157: | contig of 697 bp in length | | | |
| * 33158 | 33257: | gap of 100 bp | | | |
| * 33258 | 33951: | contig of 694 bp in length | | | |
| * 33952 | 34051: | gap of 100 bp | | | |
| * 34052 | 34735: | contig of 684 bp in length | | | |
| * 34736 | 34835: | gap of 100 bp | | | |
| * 34836 | 35483: | contig of 648 bp in length | | | |
| * 35484 | 35583: | gap of 100 bp | | | |
| * 35584 | 36308: | contig of 725 bp in length | | | |
| * 36309 | 36408: | gap of 100 bp | | | |
| * 36409 | 37108: | contig of 700 bp in length | | | |
| * 37109 | 37208: | gap of 100 bp | | | |
| * 37209 | 37893: | contig of 684 bp in length | | | |
| * 37893 | 37992: | gap of 100 bp | | | |
| * 37993 | 38672: | contig of 680 bp in length | | | |
| * 38673 | 38772: | gap of 100 bp | | | |
| * 38773 | 39475: | contig of 703 bp in length | | | |
| * 39476 | 39575: | gap of 100 bp | | | |
| * 39576 | 40290: | contig of 715 bp in length | | | |
| * 40291 | 40390: | gap of 100 bp | | | |
| * 40391 | 41087: | contig of 697 bp in length | | | |
| * 41088 | 41187: | gap of 100 bp | | | |
| * 41188 | 41891: | contig of 704 bp in length | | | |
| * 41892 | 41991: | gap of 100 bp | | | |
| * 41992 | 42679: | contig of 688 bp in length | | | |
| * 42680 | 42779: | gap of 100 bp | | | |
| * 42780 | 43500: | contig of 721 bp in length | | | |
| * 43501 | 43600: | gap of 100 bp | | | |
| * 43601 | 44295: | contig of 695 bp in length | | | |
| * 44296 | 44395: | gap of 100 bp | | | |
| * 44396 | 45088: | contig of 693 bp in length | | | |
| * 45089 | 45188: | gap of 100 bp | | | |
| * 45189 | 45880: | contig of 692 bp in length | | | |

Query Match

Best Local Similarity

Matches 254; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Score 32.98; DB 2; Length 138097;

Pred. No. 1.2e-46;

| | | |
|-----------|--|--------|
| QY 1 | ttccagtgctccaggcagcctcagcacagaaaagaacatggtctagactgaagtaccaa | 60 |
| Db 112642 | TTCCAGTGTCTCCAGCAGCCTCAGCACAAAGAAAGAACATGCTAGACTGAAGTACCAA | 112583 |
| QY 61 | ctaatcatctcttcaaatattaccgcagacacatcatggtattcaagcaccgacacag | 120 |
| Db 112582 | CTAAATCATCTCTTTCAATATTACCGCACCATCATGATTCAAGCACCGCACACAG | 112523 |
| QY 121 | tcgggtgttctgtattctctccagaatactcttcagaataatggtccacagaact | 180 |
| Db 112522 | TCCGGTGTTCCTGGTATTCTCTCCAGAAATCACTCTTCAGAAATGATGTCACAGAACT | 112463 |
| QY 181 | ttcagccagcagccttttcaactcaagcccttgcacaaattatttgcataaaaaatgaa | 240 |
| Db 112462 | TTCAGCCAGCAGCCTTTTCAACTCAAAGCCCTTGCAAAATATTTTGTCTAGAAAAATGAA | 112403 |
| QY 241 | aatcttagggactatccaga | 260 |
| Db 112402 | AATCTTAGGGTAAGTAAGA | 112383 |

RESULT 12

AC015840

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC015840

Homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered pieces.

AC015840

AC015840.2

GI:7657730

HTG; HTGS_PHASE1; HTGS_DRAFT.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 147788)

Homo sapiens, clone RP11-24D1

2 (bases 1 to 147788)

Unpublished

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Bouckhagalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A., Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donellan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karates, A., Klein, J., Lehotzky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vasiliiev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced gi:6446828.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L4434
Center clone name: 24_p1

----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 136347 bases at least Q40
Consensus quality: 143600 bases at least Q30
Consensus quality: 146212 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 147388; sum-of-contigs
Quality coverage: 5.2 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved. 6594: contig of 6594 bp in length
* 1
* 6595 6594: gap of 100 bp
* 6695 16187: contig of 9493 bp in length
* 16188 16287: gap of 100 bp
* 16288 25089: contig of 8802 bp in length
* 25090 25189: gap of 100 bp
* 25190 63923: contig of 38734 bp in length
* 63924 64023: gap of 100 bp
* 64024 147788: contig of 83765 bp in length.

FEATURES
source
Location/Qualifiers
1..147788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-24D1"
/clone_lib="RPC1-11 Human Male BAC"
1..6594
/note="assembly_fragment
clone_end:77
vector_side:left"
6695..16187
/note="assembly_fragment"
16288..25089
/note="assembly_fragment
clone_end:SP6
vector_side:right"
25190..63923
/note="assembly_fragment"
64024..147788
/note="assembly_fragment"
45466 a 29482 c 28861 g 43569 t 410 others

BASE COUNT
ORIGIN

Query Match 32.9%; Score 250.4; DB 2; Length 147788;
Best Local Similarity 97.7%; Pred. No. 1.2e-46;
Matches 254; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ttccagtgtctccaggcagcctcagcacaagaagaacatggtcttagactgaagtaccaa 60
|||||
Db 127950 TTCCAGTGTCTCCAGGCGCCTCAGCACAAGAAAGAAATGCTCTAGACTGAAGTACCAA 128009
|||||

Qy 61 ctaaatcatctcttcaaatattatcacgacacaccatcattgattcaagccgacacag 120
|||||
Db 128010 CTAATCATCTCTCTCAAAATTATCACCGACACCATCATGTTCAAGCCGACACAG 128069
|||||

Qy 121 tccggtgttcttggtatttctctcagaataatcactgtcttcagaataatggtccacagaact 180
|||||
Db 128070 TCCGGTGTCTTGTTCTCTCCAGAAATCATCTGCTTCAGAAATATGATCCACAGAACT 128129
|||||

Qy 181 ttacgacacagaccttttcaactcaaaagcccttgcaaaaattatttgcgtagaaaaatgaa 240
|||||
Db 128130 TTCAGCCACGACCTTTTCAACTCAAAAGCCCTTGCAAAAATTTATTGCTAGAAAAATCAA 128189
|||||

Qy 241 aatcttagggactactcaga 260
|||||

Db 128190 AATCTTAGGGTAAAGTAAGA 128209

RESULT 13
AP003127
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-710G6, complete
DEFINITION sequences.
ACCESSION AP003127
VERSION AP003127.2 GI:17939962
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-710G6.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161238)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsr.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Dec 19, 2001 this sequence version replaced gi:12597183.
FEATURES
source
Location/Qualifiers
1..161238
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-710G6"
BASE COUNT 50087 a 31927 c 31564 g 47660 t
ORIGIN

Query Match 32.9%; Score 250.4; DB 9; Length 161238;
Best Local Similarity 97.7%; Pred. No. 1.1e-46;
Matches 254; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ttccagtgtctccaggcagcctcagcacaagaagaacatggtcttagactgaagtaccaa 60
|||||
Db 133499 TTCCAGTGTCTCCAGGCGCCTCAGCACAAGAAAGAAATGCTCTAGACTGAAGTACCAA 133558
|||||

Qy 61 ctaaatcatctcttcaaatattatcacgacacaccatcattgattcaagccgacacag 120
|||||
Db 133559 CTAATCATCTCTCTCAAAATTATCACCGACACCATCATGTTCAAGCCGACACAG 133618
|||||

Qy 121 tccggtgttcttggtatttctctcagaataatcactgtcttcagaataatggtccacagaact 180
|||||

```

* the accession number will be preserved.
*
FEATURES             source
Location/Qualifiers
1..166793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-729B4"
/clone_lib="RPCI-11 Human Male BAC"
BASE COUNT          46459 a 31834 c 32914 g 55586 t
ORIGIN
Query Match          32.9%   Score 250.4;   DB 2;   Length 166793;
Best Local Similarity 97.7%   Pred NO 1.1e-46;
Matches 254;   Conservative 0;   Mismatches 6;   Indels 0;   Gaps 0;

QY 1 ttccagtgctccagcgagcgtccagcaagaagaacatggtctagctagtgagtcacaa 60
Db 144170 TTCCAGTGTCTCCAGCGACCTTCAGCAACAAGAAAGACATGGTCTAGACTGAAGTACCAA 144111

QY 61 ctacatcatctctctcaaatattaccgcacacatcatggtatcaagcaccgcacacag 120
Db 144110 CTAATCATCTCCCTTCAAAATATCACCGACACCATCATGGATTCAAGCACCGCACACAG 144051

QY 121 tccggtgttctgtgatttctccgaataatcatgtgttcagaataatgagtcacagaact 180
Db 144050 TCCGGTGTGTTCTGTGATTTCCTCAGAAATCACTGCTTCAGAAATATGAGTCCACAGAACT 143991

QY 181 ttacgcccagcacttttcaactcaagccctctgcagaaatattttgctagaaaaatgaa 240
Db 143990 TTCAGCCAGGACCTTTTCACTCAAGACCCCTTGCAAAATATTGCTAGAAAAATGAA 143931

QY 241 aatcttagggactatccaga 260
|||||
Db 143930 AATCTTAGGGGTAAGTAAGA 143911

RESULT 15
AP001034
LOCUS              167934 bp   DNA   linear   PRI 06-NOV-2001
DEFINITION         Homo sapiens genomic DNA, chromosome 11q clone:RP11-729B4, complete
                     sequence.
ACCESSION           AP001034
VERSION             AP001034.5   GI:16751479
KEYWORDS            HTG.
SOURCE              Homo sapiens
ORGANISM            Homo sapiens
REFERENCE
AUTHORS             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 167934)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suhiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@isc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 5, 2001 this sequence version replaced gi:11176992.
COMMENT
FEATURES
source
1..167934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-729B4"
BASE COUNT          56061 a 33108 c 32009 g 46756 t
ORIGIN
Query Match          32.9%   Score 250.4;   DB 9;   Length 167934;

```

Best Local Similarity 97.7%; Pred. No. 1.le-46;
Matches 254; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ttccagtgctccaggcagcctcagcacagaaagaaacatggtcttagactgaagtaccaa 60
Db 22624 TTCCAGTGCTCCAGGCAGCCTCAGCACAGAAAGAAAGACATGGCTCTAGACTGAAGTACCAA 22683
Qy 61 ctaaatcatctcctttcaaatattaccgcagcacaccatcatggattccaagcaccgcacacag 120
Db 22684 CTAAATCATCTCCTTTCAAAATTATCACCGACACCATCATGGATTCAAGCACCGCACACAG 22743
Qy 121 tccgggtgttctggtatttctctccagaabaatcactgcttcagaatataggtccacagaact 180
Db 22744 TCCGGTGTGTGCTGATTTCTCCAGAAATCAGTCTTCCAGAAATATGATCCACAGAACT 22803
Qy 181 ttcagccacagaccttttcaactcaagcccttgcacaaattatttgcagaaaaatgaa 240
Db 22804 TTCAGCCACGACCTTTTCAACTCAAGCCCTTGCAAAATATTTTGTAGAAAAATGAA 22863
Qy 241 aatcttagggactatccaga 260
Db 22864 AATCTTAGGGGTAAGTAAGA 22883

Search completed: August 30, 2002, 02:46:47
Job time: 7480 sec

| | | | |
|----|-----|---|-----|
| Qy | 350 | tcagatcatcattctgggctctgtttgttcatttaattctggagccttcctaattgca | 409 |
| Db | 388 | TCAGGATATCCATTCTGGGGCTCTGTTTGTTCATTAATTTGGAGCCCTCCATAATTGCA | 329 |
| Qy | 410 | gtgaagaagaaacacagagaactctgataattatgaagccgaataatgaatttcttagt | 469 |
| Db | 328 | GTGAAGAAGAAACCAAGAAACTCTGATATATATTGAGCCGAATATGAAATTTCTTAGT | 269 |
| Qy | 470 | gacctgggagcaatagctgggaatcattctcctcacatttggtttcctcctagatcaaaac | 529 |
| Db | 268 | GCCTGGGAGCAATAGCTGGAATCATTTCTCTCACATTTGTTTCATCTAGATCAAAAC | 209 |
| Qy | 530 | tacatttgggtattctcaccaaaaagtcgaagtgaagcgttactgtctctgttcttctg | 589 |
| Db | 208 | TACATTTGGTATTATTCACCAAAATAGTCAGTGTAAAGCTGTTATGTGCTCTGTTCTT | 149 |
| Qy | 590 | ggaatttgcattacattgatgacttcacgaattatgaattatcatttctctgccttc | 649 |
| Db | 148 | GGAAATTTGATTACATTTGATGATGATTTTTCAGCATTAATTTGAAATTTCTGCTTC | 89 |
| Qy | 650 | tcaatttgggctgcactcagagagattgtgattgtgacaatgttgtgactagcactg | 709 |
| Db | 88 | TCAATTTTGGGTGCCACTCAGAGGATGTGATTGTGAACAATGTTGTTGACTAGCAGCTG | 29 |
| Qy | 710 | tgagaataaagatgtgttaaaatctcaa | 737 |
| Db | 28 | TGAGATAAAGATGTGTAAATAAATAAAAA | 1 |

RESULT

4

AA416972/c

LOCUS

DEFINITION

3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ; mRNA sequence.

ACCESSION

AA416972

VERSION

AA416972.1

KEYWORDS

GI:2077080

SOURCE

EST.

ORGANISM

human.

REFERENCE

1 (bases 1 to 387)

AUTHORS

Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE

WashU-NCI human EST project

JOURNAL

Unpublished (1997)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 656 Std Error: 0.00
Seq primer: -41mi3 fwd. ET from Amersham.

FEATURES

source

1..387

/organism="Homo sapiens"

/db_xref="GDB:5926570"

/db_xref="taxon:9606"

/clone="IMAGE:730041"

/clone_lib="Soares_testis_NHT"

/sex="male"

/lab_host="DH10B"

/note="Vector: pYT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from clontech laboratories, Inc., and primed with a Not I - oligo(dT) primer [5',

4

.
